# DESIGN AND CHARACTERIZATION OF HETEROCHIRAL STRAND DISPLACEMENT REACTIONS

A Dissertation

by

BRIAN E. YOUNG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

| | |
|---|---|
| Chair of Committee, | Jonathan Sczepanski |
| Committee Members, | Jean-Phillipe Pellois |
| | David Barondeau |
| | Wenshe Liu |
| Head of Department, | Simon North |

December 2019

Major Subject: Chemistry

# ABSTRACT

The nature of Watson – Crick base pairing has enabled the rational design of complex and dynamic DNA/RNA-based molecular circuits capable of detecting nucleic acids in a sequence dependent fashion *in vitro*. Given the ease by which DNA can be programmed to interact with living systems, DNA-based molecular circuits provide an attractive avenue for the sequence-specific detection of RNA biomarkers in live cells. However, the stability of exogenous nucleic acids in biological environments remains a major concern. In order to overcome this limitation, modifications to the ribose backbone or within the phosphodiester bond have been employed to increase the resistance of DNA probes to nucleolytic degradation. Most DNA modifications in routine use alter the thermodynamic and kinetic properties of DNA/RNA hybridization, making it difficult to design complex reaction networks that function in the cellular environment. L-DNA, the mirror image (i.e. enantiomer) of natural D-DNA, represents a critically underexplored modification for this application. L-DNAs have the same physical and chemical properties as their natural counterparts, but they are essentially 'invisible' to the stereospecific environment of biology. However, L-DNA cannot form contiguous WC base pairs with D-DNA/RNA, severely limiting its use in the development of sequence-specific probes for cellular nucleic acids (summarized in Chapter 1).

Chapter 2 focuses on two potential solutions to this problem, both built on the well understood rules of DNA strand displacement reactions. We report a novel toehold-mediated strand displacement reaction utilizing achiral peptide nucleic acid (PNA)/L-

DNA duplexes and demonstrate the sequence-specific recognition of D-DNA and D-RNA inputs. An alternative strand displacement design is also reported whereby chimeric D-DNA and L-DNA duplexes are designed such that recognition of a D-DNA or D-RNA input causes the concomitant melting and release of an L-DNA output. The work presented in this section represents first of their kind heterochiral strand displacements.

Following these developments, we demonstrate the stability of these heterochiral strand displacements in living cells. Direct comparisons are made to D-DNA components, as well as components containing the common 2′-O-methyl modification. This section underscores the potential stability of L-DNA circuits, as well as the 'plug-and-play' utility of adapting D-DNA circuit designs to L-DNA.

Finally, we design a model system in chapter 4 to thoroughly characterize heterochiral strand displacement. We show that strand displacements using PNA are generally slower than their all-DNA counterparts, and that in every case strand displacement occurs slower when the input and output chirality are not matched. Interestingly, this heterochiral barrier to strand displacement enhances the mismatch discrimination of heterochiral strand displacement systems.

Overall, this work identifies dynamic molecular systems capable of sequence-specifically recognizing a D-DNA input and generating an L-DNA output. These heterochiral strand displacements are fast, biostable, and lay the foundation for the design of computational DNA systems capable of identifying endogenous nucleic acids inside live cells.

# DEDICATION

I would like to dedicate this work to my parents. Without their guidance and support none of this would have been possible.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Sczepanski, and my committee members, Dr. Pellois, Dr. Liu, and Dr. Barondeau for their guidance and support throughout the course of this research.

Many thanks to my colleagues in the Sczepanski Lab, specifically Nandini Kundu, Adam Kabza, Charles Deckard, Dr. Sougata Dey and Dr. Deb Banerjee. My friends in lab were always around to listen to me ramble about whatever crazy ideas I had and provided excellent discussions on a variety of research topics and life in general. I would also like to extend my gratitude to my friends in Dr. Pellois' lab, who taught me how to do flow cytometry and were always willing to help when I inevitably had questions about cell culture. And of course, an extra thank you to the excellent faculty and staff in the chemistry department for all the work they do for us in the graduate program.

Finally, thank you to all my friends and family back home. Specifically, my parents Evelyn and Ellis Young, and my friends Remo Gonzalez and Anthony Ciavarelli. Their continued support was truly invaluable in my success here at Texas A&M.

# CONTRIBUTORS AND FUNDING SOURCES

## Contributors

This work was supervised by a dissertation committee consisting of Professors Jonathan Sczepanski (chair), David Barondeau and Wenshe liu of the Department of Chemistry and Professor Jean-Phillipe Pellois of the Department of Biochemistry and Biophysics.

The data depicted in Chapter IV was collected in part by Nandini Kundu of the Department of Chemistry.

All other work conducted for this dissertation was completed by the student independently.

## Funding Sources

Graduate study was supported by the Cancer Prevention and Research Institute of Texas (CPRIT) under Grant Number RR150038, as well as the National Institutes of Health (NIH) under Grant Numbers R35GM124974 and R21EB027855. The contents of this dissertation are solely the responsibility of the author and do not necessarily represent the official views of CPRIT or the NIH.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

# CHAPTER I

# INTRODUCTION

## 1.1 Fundamentals of nucleic acids

Nucleic acids play a central role in the operation of all biological systems. Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are biomaterials that represent two of the cornerstones of the central dogma of biology[1]. These polymers are built from a small set of monomers, called (deoxy)ribonucleotides, and they participate in a vast array of molecular processes. The canonical function of DNA is as a molecular storage system, where it encodes information for all the molecular features of life in its quintessential double-helical structure. From this genetic template, RNA is transcribed as a single-stranded entity responsible for carrying the genetic information to the ribosome for polypeptide synthesis. These polypeptide chains undergo rearrangement, either spontaneously or directed by chaperone proteins, sometimes followed by post-translational modification, to form enzymes capable of carrying out the complex molecular reactions required for life as we know it. Broadly, while the roles of DNA and protein have for the most part remained constant over the years our understanding of RNA and its place in biology has evolved considerably[2]. It is now well known that a large amount of the DNA in our genome is non-coding, i.e. it does not code for any particular protein, but these non-coding regions can still transcribe RNA with incredibly important functions[3]. Non-coding RNAs will be discussed in more detail in section 1.3.

## 1.2 DNA and RNA structure

Both DNA and RNA oligomers are biopolymers where each monomer is linked to another by a phosphodiester backbone. A 2′-deoxyribose (in DNA) or ribose (in RNA) nucleoside linked to an aromatic nitrogenous base makes up each monomeric unit. In both cases the monomers consist of purines – adenine and guanine, and pyrimidines – cytosine and thymine (DNA)/uracil (RNA) (Figure I-1). These monomers are linked such that each phosphodiester linkage connects the 5′ carbon of one monomer to the 3′ carbon of the adjacent monomer. Each base participates in well-defined interactions through hydrogen bonding, where adenine specifically recognizes thymine (or uracil in RNA), and guanine specifically recognizes cytosine (Figure I-2 A). These base pairing interactions occur between two strands, or within the same strand, in a 5′ to 3′ orientation (Figure I-2 B). It is this set of interactions that forms the basis of our genetic code, as well as the foundation for most of the interactions between nucleic acids across biology.

DNA usually consists of two polynucleotide chains that twist around each other to form the canonical right-handed double helix. The bases project inwards, between the phosphodiester backbone of each chain, and are held together by a combination of hydrogen bonding interactions and base stacking effects. The strength of these inter-strand interactions depends on a few factors: the identity of the bases participating in hydrogen bonding interactions (a G – C base pair is stronger than an A – T base pair), and the salt conditions of the buffer. Since both DNA strands have a negatively charged backbone, increasing the concentration of

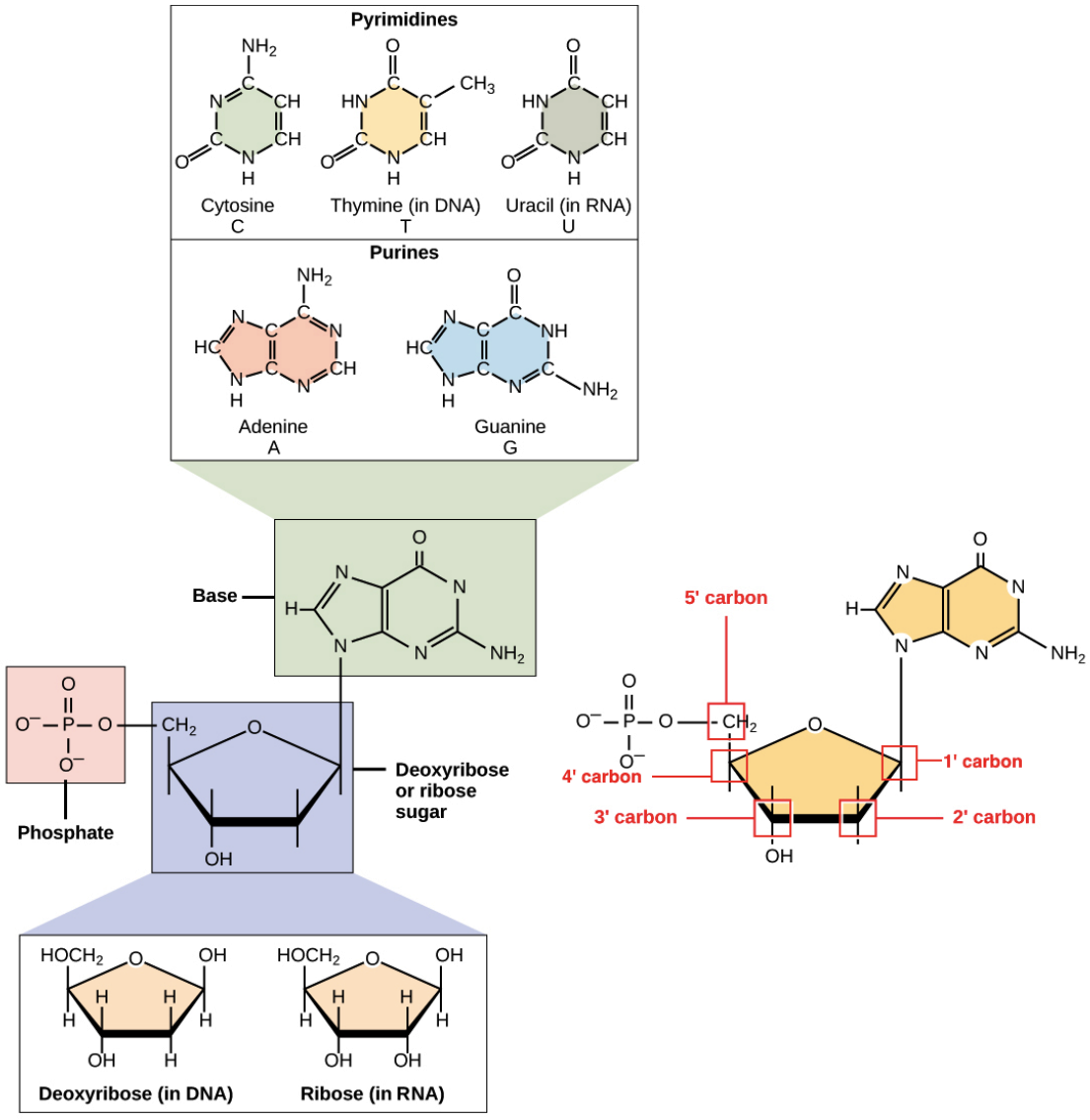



**Figure I-1. Structure of DNA and RNA monomers:** DNA and RNA monomers are composed of a deoxyribose or ribose backbone, respectively, with a 5′ phosphate and a 1′ glycosidic bond to a nitrogenous base. Adapted under CC BY 4.0, access for free at https://openstax.org/books/biology-2e/pages/1-introduction.

mono- or divalent cations can mitigate repulsive effects and increase duplex stability[4]. The DNA duplex adopts a B-form duplex with a right-handed helical turn, although particular salt conditions can promote the rearrangement of certain sequences into a non-traditional left-handed Z-form duplex (Figure I-3 A and B)[5].

RNA differs from DNA in three main aspects: the backbone of RNA contains ribose rather than deoxyribose, RNA uses the nitrogenous base uracil instead of thymine, and finally RNA is generally found in the single strand form although it has a high propensity for forming intra-strand secondary structure[6-9]. Furthermore, RNA duplexes and double stranded regions of RNA adopt an A-form duplex with a right-handed helical turn (Figure I-3 C).

### 1.3 MicroRNA, messenger RNA and RNA structure

Outside of a select few viruses, DNA is naturally found in its canonical double-helical form. Upon transcription, this genetic information is converted from DNA to RNA. There are a wide variety of RNAs that have been identified with unique functions: messenger RNA (mRNA), transfer RNA (tRNA), microRNA (miRNA), small nucleolar RNA (snoRNA) and short interfering RNA (siRNA) to name a few[10]. Many of these RNAs are known to adopt a range of complex secondary structures, such as hairpins or bulges, in addition to containing unstructured single-stranded regions. This section will focus on the structure and function of mRNA and miRNA in particular.

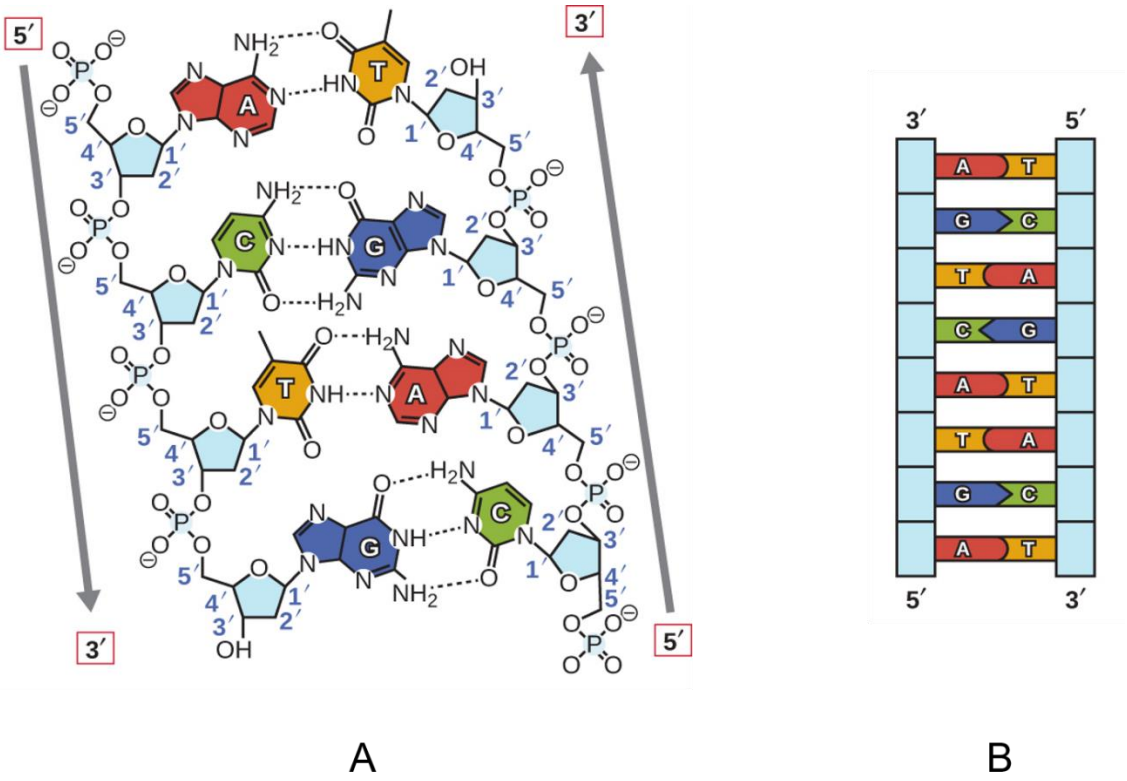



**Figure I-2. Hydrogen bonding and duplex formation.** Watson – Crick (WC) base pairing is the process whereby one DNA or RNA strand recognizes its partner via specific hydrogen bonding interactions between the nitrogenous bases. This non-covalent interaction is very specific, with deoxyadenosine recognizing deoxythymidine and forming two hydrogen bonds, and deoxyguanosine recognizing deoxycytidine and forming three hydrogen bonds. Strand recognition always occurs in the antiparallel orientation, with the 5′ end of one strand adjacent to the 3′ end of its partner. Adapted under CC BY 4.0, access for free at https://openstax.org/books/microbiology/pages/1-introduction.

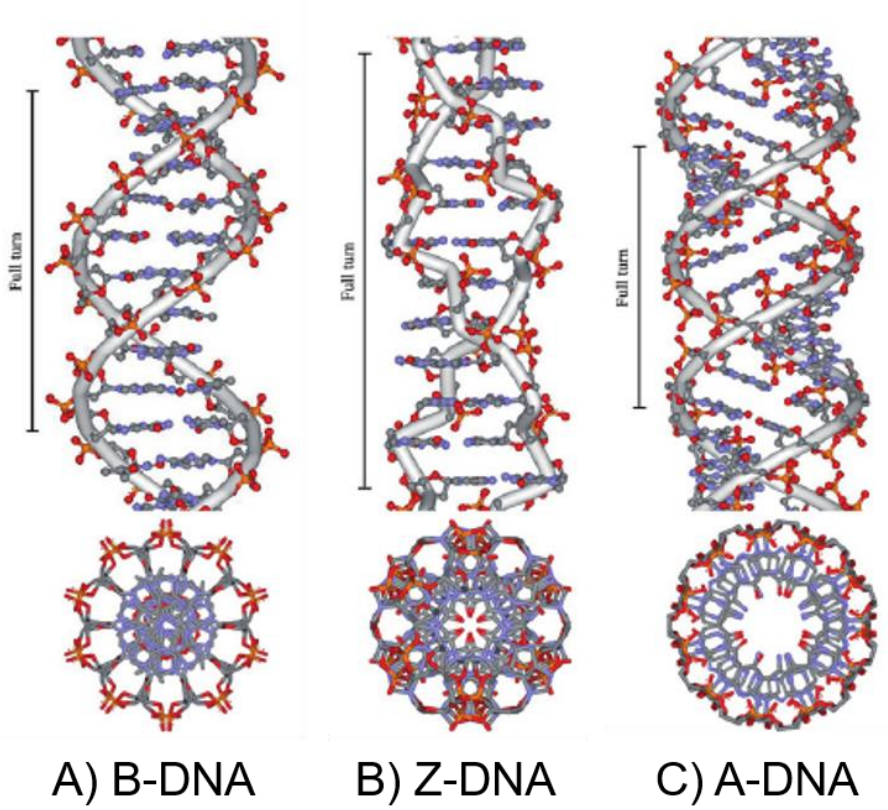



**Figure I-3. Helical conformations of DNA.** (A) B-DNA represents the canonical form of the right-handed DNA double helix. (B) Z-DNA is an alternative left-handed double helix that can be adopted under specific salt conditions. (C) A-DNA forms a right-handed double helix with geometry slightly different than that of the B-form. This A-form helix is the preferred conformation of RNA. Adapted from ref. 5 under CC BY-NC-3.0.

*1.3.1 Messenger RNA*

Messenger RNA is one of the most well studied forms of RNA and is the template from which proteins are synthesized. During transcription in the nucleus, precursor mRNAs undergo a variety of modifications before being exported to the cytoplasm[11]. It is estimated that >90% of genes in the human genome contain some level of alternative splicing pathways allowing for a far greater amount of proteomic diversity than originally expected[12-13]. This diversity is reflected in the mRNA sequence, with alternative splicing sites introducing unique sequence junctions between mRNAs that arose from the same gene.

*1.3.2 Micro RNA*

Another example of powerful cellular RNAs are miRNAs. MiRNAs are transcribed from the genome as primary microRNAs (pri-miRs), long transcripts containing hairpin features that are processed by Drosha in the nucleus (Figure I-4). These processed hairpins, or precursor microRNAs (pre-miRs) are exported to the cytoplasm where they undergo further processing by Dicer to become mature miRNAs before subsequent loading into the RISC complex (Figure I-4)[14]. These mature microRNAs are responsible for a vast network of post-transcriptional regulation and alterations in microRNA expression have been associated with increased cell survival, increased metastasis, and other disease related phenotypes[15-19].

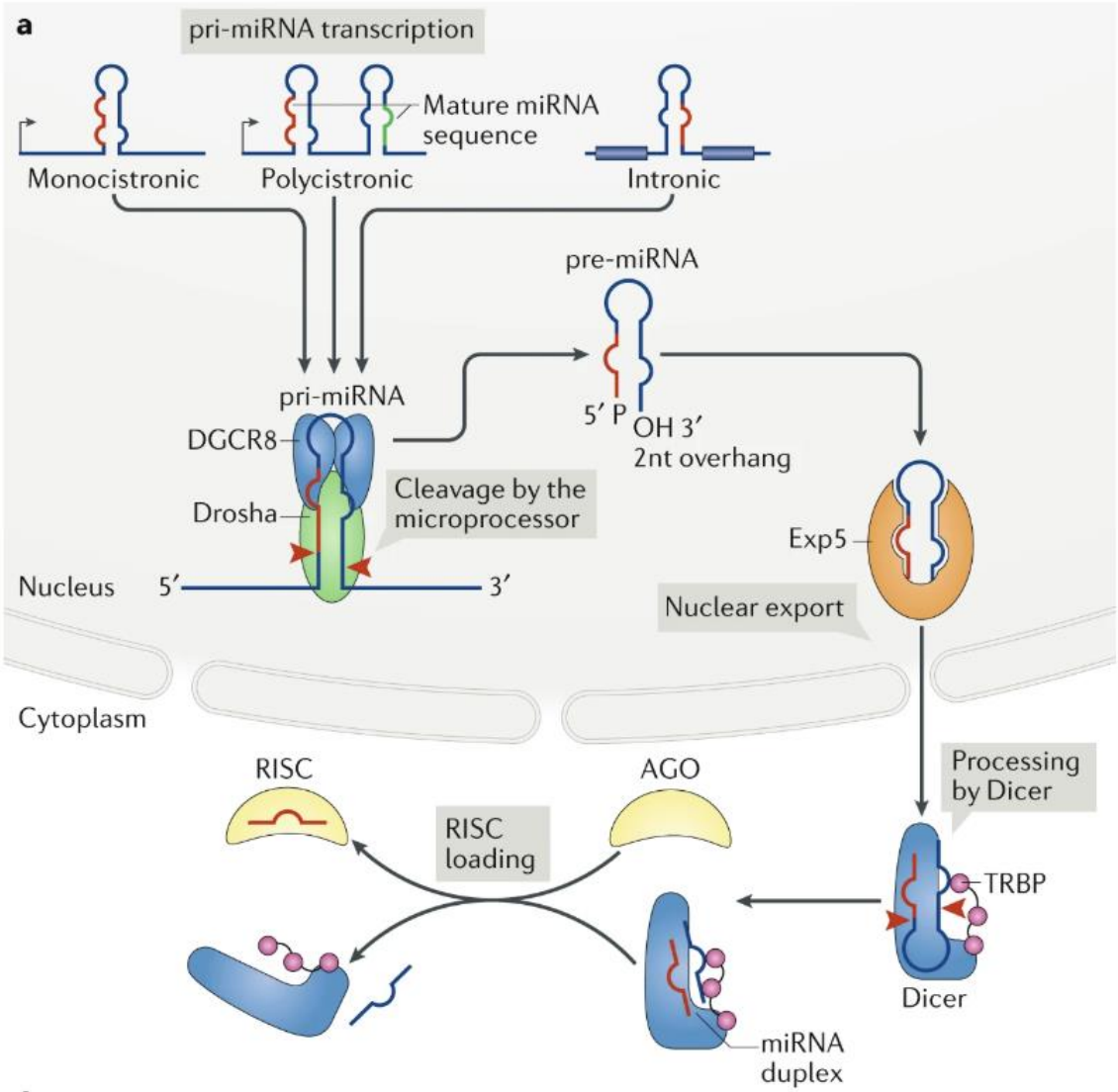



**Figure I-4 microRNA maturation.** Overview of microRNA processing and maturation. Reproduced from Ref. 14 with permission from Nature.

*1.3.3 Defects in RNA expression*

Many of today's most damaging diseases are caused by errors in our genetic code, often due to radiation or environmental damage, or defects in the DNA repair machinery. These defects also get transcribed into RNA, and they can have a profound impact on proper RNA maturation. Even simple mutations can impact the secondary structure of RNA transcripts, leading to splicing errors generating improper mRNAs[20] or microRNA strands containing point mutations that impact activity[21-22]. RNAs produced with such errors often contain unique secondary structure features, or unique sequences in single stranded regions, that provide an opportunity for their detection.

Alternatively, errors in the genetic code can increase or decrease the expression of certain RNAs[23-25]. Altered RNA expression levels are often associated with disease states, potentially leading to a decrease in cell attachment, increased resistance to cell death pathways, and increased proliferation. For example, the expression of miRNA-155 is increased in a variety of difficult to treat breast cancers, and early detection of RNA levels can increase the accuracy of disease diagnostics and help inform medical decisions[26-27].

## 1.4 Current methods for the detection of aberrant RNA expression

Due to their impact, detecting errors in RNA expression has become a driving force in both academic and industry pursuits. As such, many techniques have been developed for the detection of the overall expression of RNA transcripts, as well as sequence specific detection methods to identify improperly spliced transcripts or point mutations in mRNAs and miRNAs. In the following sections I will outline a selection of these methods, and

briefly describe some of their pros and cons with respect to RNA sensing. The techniques discussed in this section rely on the principles of WC base pairing for primer and probe design, highlighting how these rules are truly a cornerstone for the development of nucleic acid technologies.

### *1.4.1 In vitro methods*

Most conventional methods for RNA detection are performed *in vitro*, first requiring the isolation of total RNA or a target RNA from a population of cells or tissue. A non-exhaustive list is provided in table 1.1 below.

| Detection Methods | Selected References | Target RNA |
|---|---|---|
| RTqPCR | 29, 30 | mRNA, small non-coding RNAs |
| Northern Blotting | 35, 36 | mRNA, small non-coding RNAs |
| *In situ* hybridization | 44, 45 | mRNA, small non-coding RNAs |

**Table I-1.1** Common methods for the *in vitro* detection of target RNAs

#### *1.4.1.1 Real time quantitative polymerase chain reaction*

Real time quantitative polymerase chain reaction (RT-qPCR) is the current gold standard for RNA detection in the fields of molecular diagnostics, life sciences, agriculture and medicine[28]. This method requires the isolation of total RNA from cells or tissue, the removal of DNA contaminants, the production of cDNA from the RNA pool (popular

cDNA primers include: target specific primers, random primers 6-9 bases long, polyT primers that specifically amplify targets containing a polyA tail, or sometimes a careful mixture of different primers), then amplification of cDNA by carefully designed PCR primers[29-30]. This technique has also been adapted for the detection of short non-coding RNAs[31-32]. While proper use of this technique can be incredibly rewarding, with the capabilities to detect <10 RNA molecules/reaction, a number of common pitfalls can make it difficult to utilize to its full potential[33-34]. Issues such as poor RNA quality (through improper preparation, handling or storage), insufficient removal of genomic DNA or poor primer design leading to false positives, and poor validation of housekeeping genes[35] are all factors that must be taken into account when using this detection method.

*1.4.1.2 Northern blotting*

Northern blotting is an RNA detection technique similar in practice to western blotting commonly used for protein detection. As with RT-qPCR, the first step of this technique is the isolation of total RNA from the sample of interest. Depending on the target, methods can be used to further enrich the RNA isolate for specific sub-types of RNA such as mRNAs containing a polyA tail[36] and the protocols can be optimized for the detection of short RNA sequences such as microRNAs[37-38]. RNA samples are then separated by gel electrophoresis and transferred to a nylon membrane via traditional blotting techniques before being covalently immobilized on the membrane by UV light or heat. After RNAs have been captured on the membrane, the target RNA of interest is visualized by the hybridization of a complementary RNA probe generally labeled with

radioactive nuclei. Nonradioactive probes conjugated to digoxigenin with a chemiluminescent readout have been used as well, with similar sensitivity and higher throughput than traditional radioactive probes[39]. While this method is not as sensitive as RT-qPCR, it has a few advantages such as gaining information regarding the approximate size of the RNA as well as potentially learning about alternatively spliced constructs (as splicing events will affect the gel shift of the target)[36, 40].

*1.4.1.3 In situ hybridization*

*In situ* hybridization (ISH) differs slightly from the previously discussed methods in that it is used to probe specific RNAs within the subcellular environment. ISH was first described in 1969 when Gall and Pardue used radioactive RNA probes for the detection of genomic DNA in cytological preparations from *Xenopus* oocytes[41]. The technique has evolved to encompass the use of fluorescent probes (FISH) as well as *in situ* amplification techniques for higher signal generation for the detection of RNAs in a variety of sample types such as fixed cells or tissue slices[42]. Briefly, fixed cells or tissue samples are incubated with an RNA or DNA probe that is complementary to the target of interest. After sufficient time has passed to allow for target binding, subsequent washing steps remove unbound probe to increase the signal to background ratio. While there are some benefits to using radiolabeled probes, such as high sensitivity and relative ease of quantitation, these probes generally require long exposure times and tend to have poor spatial resolution[43]. Nonisotopic labeling, where probes are conjugated to small molecule dyes or antibody epitopes, can be visualized by fluorescence microscopy or

immunohistochemistry. Of these nonisotopic techniques fluorescence labeling shows great promise for the design and application of multiplexed probes, and much work has been done towards the development of *in situ* amplification of fluorescence signal[44-45]. While *in situ* hybridization and amplification techniques have the benefit of detecting RNAs within the subcellular compartments in which they reside, the technique generally has difficulty identifying low copy number targets and targets that are sequestered in protein – nucleic acid interactions. Newer protocols have increased the power of FISH techniques for detection small RNAs[46-47].

*1.4.1.4 Limitations to in vitro detection of RNA*

While these detection methods can be incredibly sensitive, with techniques like RT-qPCR being able to detect the presence of a few molecules of RNA, they generally require extensive sample handling and processing. Each additional step in the workflow adds the possibility of sample contamination that might either degrade the RNA (RNAse contamination) or generate false positive signals (genomic DNA contamination) due to such high levels of amplification. In addition to the techniques discussed in this section, massively parallel sequencing techniques like RNA-seq can provide huge amounts of data from cell populations[48]. However, this technique can still be cost prohibitive for routine use and suffers many of the same drawbacks regarding sample stability and library preparation found in other techniques in this list. The methods discussed here, with the exception of (F)ISH techniques, all detect ensemble RNA levels (the average level of RNA expression across the isolated group of cells/tissue).

*1.4.2 Live cell methods*

Detection of RNA expression levels in live cells, either in cell culture or in living organisms, is a more challenging prospect. The following techniques allow or are working towards the direct detection of nucleic acid targets in live cells without the intensive workflow and excessive sample handling requirements needed in traditional cell lysis and RNA isolation techniques, and they enable the fluorescence detection of RNAs within individual cells rather than ensemble RNA (similar to the ISH probes described in the previous section). One of the most important considerations for detecting nucleic acids in a living system is the method of delivering the probe. As such, this section will focus on some of the most common methods of probe delivery as well as the style of probes used. A non-exhaustive list of common delivery methods is provided in table 1.2 below.

| Delivery Method | Selected References | Probe Type |
|---|---|---|
| Cationic Transfection Reagents | 51, 52, 53 | HCR, CHA, SD probes |
| Electroporation | 66, 67 | MBs, SD probes |
| Nanoparticles | 68, 69 | Nanoflares, DNAzymes |
| Cell-Specific Aptamers | 70, 71 | MBs, SD probes |
| Microinjection | 57, 58 | MBs, DNAzymes |
| Cell-Penetrating Peptides | 64, 65 | CHA, MBs |

**Table I-2.** Common methods for the live cell detection of target RNAs

The most common nucleic acid probe use for the detection of RNAs in live cells is the molecular beacon[49], primarily due to their relatively straight-forward design parameters. The standard molecular beacon probe is composed of a single stranded loop region that is complementary to the target RNA, which is flanked by complementary stem sequences on either side and modified at both the 3′ and 5′ ends with fluorophore/quencher or acceptor/donor dyes. In the absence of the target the two stem regions will be hybridized, resulting in a quenched fluorescence signal (or FRET, depending on the dyes used), but once the target is available the single stranded loop region will hybridize to it. The resulting duplex has a higher Tm than the stem, making its formation thermodynamically favorable, and its formation separates the dyes at the two termini thus increasing fluorescence signal (or disrupting FRET). Due to their ubiquity this type of probe will be referenced in many of the examples below.

*1.4.2.1 Cationic Transfection reagents*

Cationic transfection reagents (TRs) are one of the most common ways to deliver exogenous nucleic acids into live cells. Cationic TRs encompass a broad range of cationic lipids (lipoplexes) and cationic polymers (polyplexes)[50]. These positively charged components interact with negatively charged DNA or RNA, forming lipoplexes (cationic lipids), polyplexes (cationic polymers), and lipopolyplexes (a mixture of cationic lipids and cationic polymers). The nanoparticles formed through this interaction can vary in size based on the +/- charge ratio, and this can have a dramatic impact on the efficiency of nucleic acid delivery[51]. Generally, cationic TR mediated delivery of DNA takes place in

two steps: 1) Binding of the complex to the cell via electrostatic interactions with the negatively charge cell membrane, where particles are mainly internalized through endocytosis and 2) Escape of DNA into the cytoplasm. Cytoplasmic escape occurs during the early endosome through interaction of the cationic TR components with the negatively charged lipids comprising the endocytic membrane, leading to a flip-flop of anionic phospholipids into the lipoplex and the generation of pores through which the DNA cargo can escape into the cytoplasm. While these methods were originally designed for the delivery of large plasmid DNA into mammalian cells[52], a number of recent publications have described the delivery of small DNA probes for the detection of cellular RNAs[53-55].

*1.4.2.2 Microinjection*

The technique for microinjecting live cells with exogenous components is over 100 years old[56], and the microinjection of nucleic acids has been described as early as 1971[57]. Microinjection has since grown in popularity as a technique to deliver genes and other functional nucleic acid molecules since it provides a straight-forward path for the introduction of these components directly into the cytoplasm. Coupled with highly specific DNA probes based on WC base pairing, this technique has been used for the real time detection of RNA sequences in live cells[58-60].

*1.4.2.3 Conjugation with cell penetrating peptides*

Cell penetrating peptides (CPPs) were first described in 1988 when it was demonstrated that the transactivator of transcription (TAT) protein from the human

immunodeficiency virus type 1 (HIV-1) was capable of efficiently crossing the cell membrane[61-62]. Future studies isolated the domain responsible for membrane translocation to a series of basic amino acids from positions 48 to 60 (sequence: RKKRRQRRR). This sequence is now known as the TAT peptide and has been used to enable the delivery of plasmid DNA[63-64] and small nucleic acid probes into the cytoplasm of live cells[65]. While there are many other CPPs, TAT remains the most well studied especially with respect to DNA delivery. Briefly, TAT has been conjugated to MB probes via a plethora of covalent linkers[66-67] and was able to deliver its cargo to cells rapidly (<30 minutes) with nearly 100% efficiency. One initial concern with this method was that the positively charged CPP might interfere with target recognition of the molecular beacon, however no peptide dependent loss of hybridization has been reported for any of the common linkage methods.

With the delivery techniques described in this section, or others such as electroporation[68-69], nanoparticle conjugation[70-71], or direct targeting by cell-specific aptamers[72-73], once inside the cell exogenous nucleic acid probes face a series of challenges impacting their ability to find and correctly identify the desired target. Delivery methods proceeding through endocytosis introduce DNA probes to a variety of nucleases capable of rapidly degrading DNA, leading to target agnostic fluorescence enhancement and high background. Even DNAs delivered directly into the cytoplasm face nuclease degradation, in addition to off target effects due to partial sequence complementarity[74] and ubiquitous DNA binding proteins[75]. Additionally, it has been demonstrated that small DNA probes delivered to the cytoplasm often end up sequestered in the nucleus where

they cannot interact with the intended target and demonstrate increased non-specific signal generation[60]. A variety of methods have been attempted to address these issues. Certain DNA modifications have been shown to increase nuclease resistance[69] as well as decrease non-specific interactions with DNA binding proteins[76], and probes have been designed with additional nuclear export sequences to decrease accumulation in the nucleus[59].

### 1.5 DNA strand displacement reactions

Synthetic DNA probes have become valuable tools in the detection of aberrantly expressed nucleic acid targets, both *in vitro* and *in vivo*. The robust simplicity of Watson – Crick base pairing allows for the rational design of DNA-based probes, and the advent of solid phase DNA synthesis techniques enables the synthesis and modification of nearly any probe regardless of sequence. I will address DNA synthesis and modification techniques in section 1.6, but this section will focus on the utility of DNA in the development of dynamic nanoscale devices.

Using DNA as a molecular switch was first realized in 1999, when Nadrian Seeman assembled an all DNA device consisting of two rigid double cross-over (DX) motifs connected by a 4.5 turn DNA "proto-Z" duplex[77]. Each DX motif was conjugated with a fluorophore (one with fluorescein and one with cyanine 3) such that the dyes were close enough to participate in föster resonance energy transfer (FRET) in the nano-device's natural state. However, the proto-Z DNA "axel" connecting the rigid DX tiles can be converted from the canonical right-handed B-form duplex to a left-handed Z-form duplex depending on the ionic strength of the solution. This transition causes the "axel"

to twist in a manner that forces the rigid DX tiles apart, resulting in a loss of FRET between the two dyes. This nanodevice could be cycled between ON and OFF states, was the first example of a man-made DNA device capable of molecular motion.

### *1.5.1 Dynamic DNA devices in vitro*

After Seeman's seminal work regarding DNA's ability to perform molecular work in response to environmental factors, Neumann's lab reported a DNA device capable of responding to a separate DNA fuel strand[78]. This so-called molecular tweezer is composed of 3 strands forming a nicked duplex with single-stranded overhangs at either end. Strand A was labeled at the 5′ end with tetrachlorofluorescein (TET) and the 3′ end with carboxy-tetramethylrhodamine (TAMRA), then hybridized to strands B and C. Strands B and C each contained a 5′ or 3′ single-stranded overhang, respectively. In this default state the fluorophores were held apart due to the semi-rigid nature of the nicked duplex, but upon addition of a DNA fuel strand (F) complementary to both domain B and domain C's single-stranded overhangs the device would be closed. This motion brings the fluorophore labels into proximity, enabling FRET and locking the tweezers in a rigid, closed state. Strand F was designed with its own 8-base single stranded region, enabling a 5$^{th}$ strand, F*, to hybridize to this single stranded domain and displace B and C through a process called branch migration[79]. Through iterative additions of strands F and F*, the tweezers could be cycled through ON and OFF states at least 7 times. This work represents the first example of a rationally designed DNA toehold being used to fuel work on nanoscale DNA devices, and the concept of

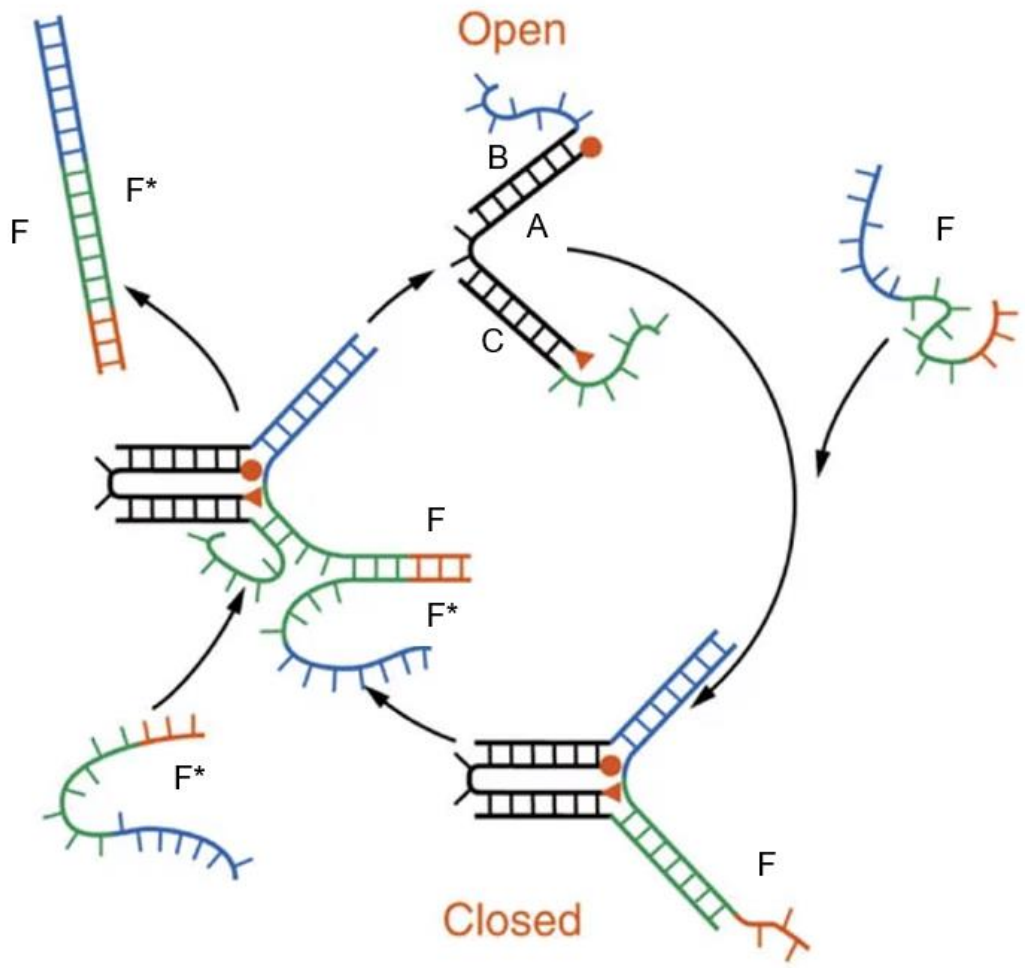



**Figure I-5. DNA driven molecular motion.** Strands A, B and C are annealed together to form the molecular tweezers. Addition of strand F, which is complementary to the single stranded portions of strands B and C (blue and green domains, respectively), closes the tweezers through hybridization and quenches fluorescence. Subsequent addition of strand F* reverses this reaction, by binding to the single stranded domain on strand F (shown in orange) and displacing this strand from the tweezers through branch migration. This displacement restores fluorescence, and the tweezers can be cycled between ON and OFF states by further additions of strands F and F*. Adapted from Ref. 62 with permission from Nature.

"toehold mediated strand displacement" has become ubiquitous within the field of DNA nanotechnology.

In 2006, Eric Winfree expanded the utility of toehold-mediated DNA strand displacement (DSD), describing a set of toehold-containing DNA systems capable of performing Boolean logic operations[80]. These DNA-based logic gates were engineered to emulate digital design principles: logic, cascading, restoration, fan-out and modularity. This work laid the foundation for complex DNA computations that can be used to detect specific nucleic acids, even specific combinations of nucleic acids, within complex reaction mixtures. Interestingly, this molecular computation approach provides an alternative to traditional multiplexing techniques which usually require multiple spectrally resolved fluorophores to accomplish the simultaneous detection of multiple independent nucleic acid sequences. Although much of Winfree's recent work has focused on scaling up these molecular computers, having recently reported systems containing >100 unique DNA strands[81], I will focus the remainder of this section on 4 important DNA motifs: The translator gate, the reporter, the AND gate, and the OR gate. A translator gate consists of two DNA strands forming a short duplex with single-stranded overhangs at either the 3′ and 5′ ends of each strand (Figure I-6, complexes A, B and C). One of these overhangs will function as the toehold, acting as a nucleation site for the desired input to bind before displacing the incumbent strand from the duplex through branch migration, while the other extends the sequence of the output strand to nearly any desired length. In this way a translator gate 'translates' the sequence of the input strand to the desired

sequence of the output strand, with no specific sequence overlap required except for what will act as the new toehold for downstream components. The reporter motif represents the digital "readout" of a DNA strand displacement cascade. After all desired logic operations have been completed, the final strand that is released will toehold invade a reporter duplex that has been modified to contain either a fluorophore/quencher pair or a donor/acceptor pair (Figure I-6, complex D). An AND gate has a similar design but is composed of 3 strands with two toehold sites (Figure I-6, complex D). In this design one of the toeholds is sequestered within the duplex region and is only revealed once the first input binds. Binding of the fist input proceeds through toehold nucleation followed by branch migration, and when the first incumbent strand is displaced the second toehold becomes single stranded. At this point the second input can nucleate and branch migrate through the second incumbent strand, completing the operation of the AND gate. AND gates can be used to directly detect desired inputs, or inputs can activate translator gates before cascading into the AND gate. An OR gate is similar in concept to an AND gate, in the sense that it is responsive to multiple inputs, but while an AND gate requires both inputs to be activated an OR gate can be activated to the same extent by either single input or both inputs together. This strand invasion will spatially separate the fluorophore from the quencher (or the acceptor from the donor), generating a signal proportional to the success of the logic operation.

These computations can be rationally designed due to the incredible specificity of WC base pairing. Thanks to decades of research, our understanding of

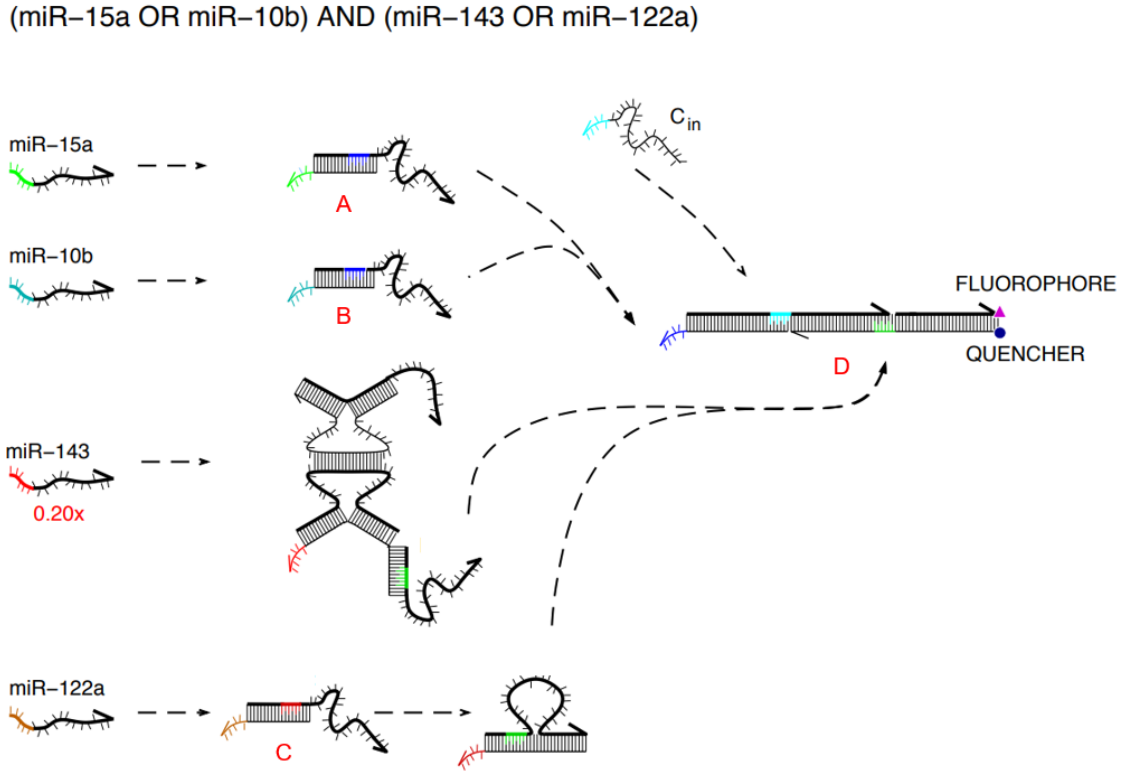



**Figure I-6. A multi-layered DNA circuit.** A four input DNA "computer," that generates fluorescence based on the presence of (miR-15a OR miR-10b) AND (miR-143 OR miR-122a). Complexes A, B and C represent translator gates that convert their respective inputs into new sequences to interact with the reporter module. Complex D is the reporter and is acting as both an AND gate and an OR gate. Either miR-15a or miR-10b activating their translator gate will enable strand $C_{in}$ to bind to the light blue toehold on the reporter and reveal the final green toehold domain. The presence of either miR-142 or miR-122a will release a new strand capable of binding to the green toehold domain and displacing the fluorophore from the quencher, activating the reporter and generating fluorescence. Adapted from reference 64 with permission from AAAS.

the thermodynamics and kinetics of DNA hybridization is near-quantitative, further enabling the computer aided design of more complex systems[82-83]. Still, slow reaction kinetics or spurious activation of circuit components are challenges that need to be addressed on a per-design basis. This is generally achieved by ensuring that all sequences not constrained due to target complementarity don't have a high risk of secondary structure formation, to an extent this factors into target selection as well, as unforeseen secondary structure can retard strand invasion or accelerate leak reactions[84]. Leak in these strand displacement systems predominately arises from fraying at the edge terminal base pairs of duplex regions (Figure) transiently revealing sequences that can be invaded in the absence of the correct target. Leak of this type can be mitigated through the use of short clamp regions composed of strong base pairs that are not complementary to downstream components[85], or by increasing the energy barrier of branch migration[86].

*1.5.2 Dynamic DNA devices in vivo*

As the capability of DNA strand displacement increased, the desire to utilize such reactions for the sequence-specific detection of nucleic acids in live cells was an apparent next step. To date, a handful of groups have demonstrated the *in vivo* detection of RNA targets with strand displacement reactions based on the designs outlined in the previous section[53, 87]. The first consideration in the design of such systems is their stability in the cellular environment. It has been reported that exogenous DNA introduced into cells is rapidly degraded by endo- and exonucleases, leading to decreased circuit performance

(toehold degradation) and increased background signal (degradation of reporter domains). Component stability is addressed by the incorporation of modified DNA residues within the system. While these modifications increase stability, they often have dramatic effects on the thermodynamics and kinetics of strand hybridization which undermines the perceived simplicity of circuits designed by Watson – Crick base pairing[88-90]. Additionally, many of these modifications have been reported to increase the toxicity and immunogenicity of exogenous DNA[91-92]. These issues represent considerable challenges towards the routine operation of DNA strand displacement circuits in live cells.

Once the system has been designed and modified appropriately, all the required components must be delivered into the cells where they have the chance to interact with their intended target. Common delivery methods for DNA-based probes were discussed briefly in section 1.4.2, and many of those methods have been used to deliver DNA circuits as well. The most common method of component delivery is through the use of cationic lipid systems such as Lipofectamine 2000/3000™. While these reagents are powerful tools for the delivery of DNA into living cells, their interactions with DNA, and the effects of those interactions on strand displacement performance are not well understood, although it has been reported that complexation of DNA components with some transfection reagents can prevent their interaction *in vitro*[87, 93]. Many traditional delivery reagents tend to result in nuclear localization of small DNA components (ref 71 and our own work), likely due to passive transport across nuclear pores, but many desirable RNA targets are found almost exclusively in the cytoplasm. To prevent this, DNA probes have been appended to bulky substituents that prevent access to the nucleus (XX), or they can be

synthesized with modified RNA tags acting as nuclear export signals (XX). Alternatively, these uncertainties can be circumvented through the direct modification of DNA with small molecules that enhance cellular uptake and delivery. DNA modification with hydrophobic small molecules such as cholesterol, α-tocopherol (Vitamin E) and stearyl residues have been shown to increase cellular uptake *in vivo*[94-95]. Such modifications are expected in increase the interaction of DNA molecules with lipoprotein particles, lipoprotein receptors, and transmembrane proteins[94]. Some more direct delivery approaches have been reported, with the most common being conjugation of DNA to one or more N-acetylgalactosamine (GalNAc) moieties. This modification directly interactions with the asialoglycoprotein receptor (ASGPR), which is highly expressed on the surface of liver cells and rapidly internalized[96].

### 1.6 Nucleic acid synthesis, modifications, and analogues

The design and synthesis of nucleic acid oligomers entered somewhat of a renaissance in the early to mid-1990s, leading to the development of automated synthesizers capable of chemically synthesizing sequence-specific, single-stranded DNA and RNA oligonucleotides. A number of synthetic strategies were evaluated[97-99], eventually leading to the development of phosphoramidite building blocks by the Caruthers lab in the early 1980s[100-101] (Figure 1.6.1). These monomers are relatively stable solids that can be stored for extended periods of time prior to use and exhibit rapid coupling times with greater than 98.5% routine coupling yields. ON synthesis is a solid

phase synthesis technique, with the most common supports being functionalized controlled-pore glass (CPG) or polystyrene (PS).

This iterative synthesis cycle enables the site-specific incorporation of any modification that is amenable to the chemistry behind the phosphoramidite approach. A selection of these modifications, and their common applications, will be discussed in the sections below.

### *1.6.1 Common ON modifications*

Many of the most common ON probe modifications are employed to increase their resistance to nuclease degradation. This is most often achieved by modification of the ON backbone, such as phosphorothioate (PS) and methylphosphonate linkages, or by modification of the 2′ position of the deoxyribose sugar. Common 2′ modifications include fluorine (2′-F), O-methoxy (2′-OMe), and O-methoxyethyl (2′-MOE), all of which show a moderate to large increase in nuclease resistance while slightly increasing the Tm of duplex formation. While these are desirable properties, some of these modifications have been shown to increase toxicity and immunogenicity *in vivo* which may limit their potential uses.

There are many non-base modifications that are routinely used in DNA synthesis. Modifying DNA strands with fluorophores enabled the development of *in situ* imaging of fixed or live cells, and quencher modifications are routinely used in the design of ‘digital’ outputs in DNA strand displacements and computations. Hydrophobic modifications such as those discussed in section 1.5.2 can be routinely incorporated as well. The availability

of phosphoramidite modifications enables the synthesis of a huge amount of unique DNA sequences and designs, and the precision of WC base pairing allows us to rationally design complex structures and dynamic systems from these chemical oligomers.

### *1.6.2 L-DNA as an ON modification*

L-DNA will be the primary modification discussed in the following chapters. This is an interesting backbone modification in which the enantiomer of the naturally occurring D-(deoxy)ribose, L-(deoxy)ribose, is incorporated into the backbone of oligonucleotides (ONs). These enantiomeric building blocks have been prepared as phosphoramidites and as such the design and synthesis of ON oligomers containing an L-ribose backbone is straightforward. As enantiomers, L and D-ONs have identical physical properties in terms of solubility, hybridization kinetics, and duplex thermal stability[102-103]. This means that any design principles established for D-ON-based systems can be applied directly to L-ONs. This dramatically simplifies the optimization steps when transitioning from unmodified to modified systems as compared to other common modifications. However, due to the incredible stereospecificity found in nature, enzymes cannot recognize L-nucleoside monomers or L-oligomers, and as such DNA and RNA strands with an L-(deoxy)ribose backbone can only be synthesized chemically. By this same property, L-ONs are resistant to nuclease degradation, nontoxic, and non-immunogenic[102, 104-105]. Together, these characteristics make L-ONs an attractive option for applications in biological environments. However, L- and D-ONs are incapable of forming contiguous WC base pairs with each other[103, 106]. This property prevents the rational design of L-ON

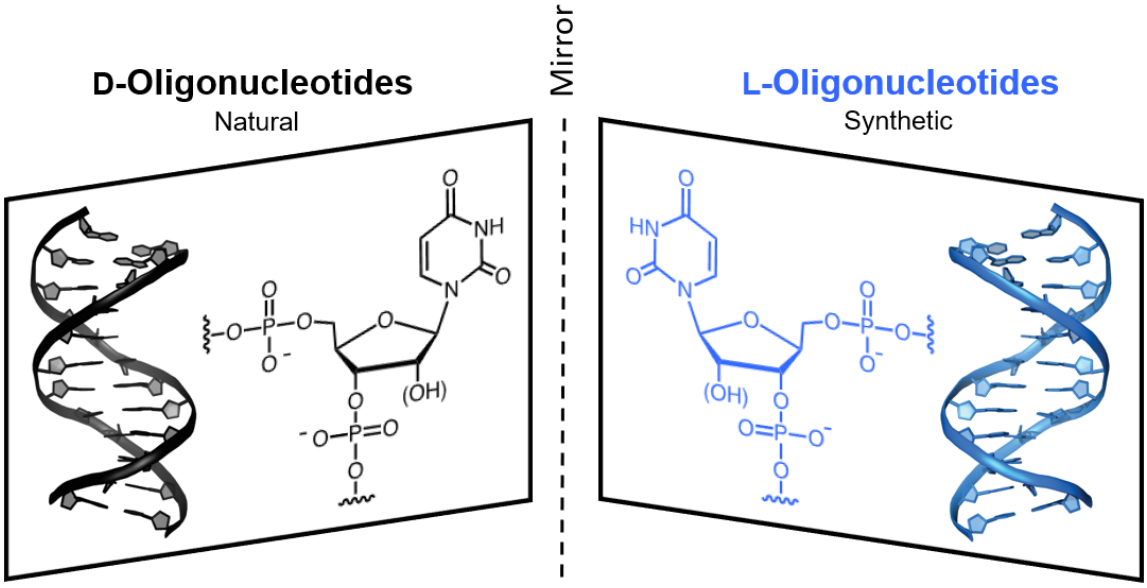



**Figure I-7. A comparison of D-DNA and L-DNA.** D-DNA and D-RNA, shown on the left in black, is the nature chirality of these oligonucleotides. L-ONs, shown on the right in blue, are the synthetic enantiomers of their natural counterparts.

hybridization-based tools for the detection of native D-DNA or D-RNA targets, one of the key design parameters in the development of *in vivo* DNA probes and computational systems.

*1.6.3 Peptide Nucleic Acid*

Peptide nucleic acids (PNAs) are more than a simple DNA modification, these molecules are synthetic analogues of DNA (Figure I-8). PNA contains the same nitrogenous bases as DNA and RNA, but they are attached to an achiral N-(2-aminoethyl)glycine backbone instead of the traditional D-(deoxy)ribose backbone[107-108]. This analogue was first synthesized by Nielsen and colleagues in 1991 is are resistant to nuclease and protease degradation. Due to the nature of the PNA backbone, this polymer is achiral and was found to have a higher $T_m$ for complementary nucleic acid targets (both DNA and RNA) as compared to traditional DNA/DNA, DNA/RNA and RNA/RNA hybrids[107, 109]. This increased $T_m$ is likely due to the neutral PNA backbone, avoiding electrostatic repulsion with the negatively charged backbone of its DNA and RNA complement. As such, the binding affinity of PNA to DNA or RNA is independent of salt concentration.

PNA is capable of hybridizing to its complement in the traditional anti-parallel orientation, as well as the parallel orientation[110-111]. While the "correct" antiparallel orientation is strongly favored ($\Delta T_m$ ~13 ºC for the pentadecamer PNA H-TGTACGTCACAACTA-$NH_2$), it is important to note that the $T_m$s of parallel binding are similar in strength to canonical DNA/DNA, DNA/RNA, and RNA/RNA duplexes. PNA/DNA heteroduplexes adopt a duplex structure somewhat similar to that of a

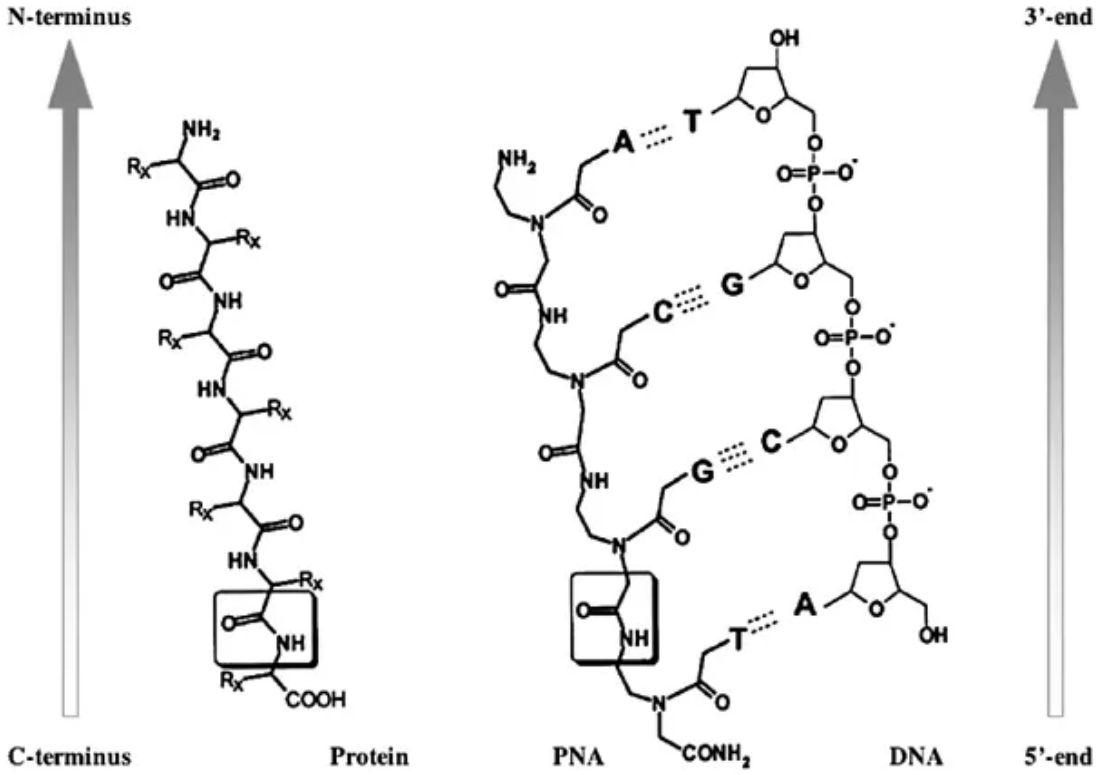



**Figure I-8. Peptide nucleic acids and DNA.** PNAs are synthetic analogues of DNA built with a 2-aminoethyl glycine backbone similar to the amide backbone in proteins. PNA is achiral and hybridizes to DNA and RNA through WC base pairing. Reproduced from Ref. 96 with permission from Nature.

DNA/DNA duplex, termed a P duplex, with both A- and B-form character[112]. PNA/RNA heteroduplexes adopt an A-form duplex, similar to that of RNA/RNA duplexes[113]. The kinetics of PNA/DNA and PNA/RNA hybridization are altered as well, with these hybrid duplexes displaying faster ON rates ($k_{on}$) and slower OFF rates ($k_{off}$) than their natural counterparts[109].

While PNA has many potentially useful properties, such as strong affinity for both DNA and RNA targets, and high resistance to enzymatic degradation, and a higher propensity for mismatch discrimination it suffers a few drawbacks. Since the PNA backbone isn't charged, it is generally more hydrophobic than its same-sequence nucleic acid counterpart. Thus, as the PNA length increases, so too does the risk of aggregation and insolubility[114]. This can be overcome by the modification of PNA with PEG-like linkers or positively charged amino acid residues but depending on the desired application this may not be ideal. As such, PNA design elements such as length and purine:pyrimidine base ratio are incredibly important when choosing a PNA sequence for an intended application. Additionally, PNA can form incredibly strong base pairing interactions with other PNA sequences, dramatically increasing the risk of self-complementarity within PNA strands as well as the risk of using multiple PNAs within the same system.

### 1.7 Goal of this research

Dynamic DNA nanotechnologies based on toehold mediated strand displacement have become powerful tools for the design of *in vitro* computational systems and *in vivo* sensors for the detection of nucleic acids due to the utility of WC base pairing. However,

the propensity for the degradation of exogenous DNA probes, as well as their potential for off target interactions with other biomacromolecules, has limited the application of DNA – based systems for live cell imaging to relatively simple designs. While a near-ideal DNA modification exists to address these problems in the form of L-DNA, the inability of this biopolymer to form contiguous WC base pairs with natural D-DNA or D-RNA eliminated its utility in the development of sequence specific sensors. The work outlined in this dissertation begins to address this shortcoming using achiral PNAs and chimeric D/L-ONs in the design and characterization of heterochiral strand displacement systems. For the first time, these approaches have enabled the displacement of biostable L-DNA strands dependent on the sequence specific recognition of a D-DNA or D-RNA input. With these novel strand displacement systems in hand, we wanted to address two key questions: Will strand displacement components made of L-DNA remain stable and function as intended in live cells? And how do these novel PNA/L-DNA reactions compare to the current body of all-DNA strand displacement reactions in terms of specificity and kinetics? The results of this work will help inform the design of next generation circuit components, and hopefully enable the development of higher-order DNA reactions for the detection of endogenous nucleic acids in live cells.

## 1.8 References


1. Crick, F., Central Dogma of Molecular Biology. *Nature* **1970,** *227* (5258), 561-563.

2. Licatalosi, D. D.; Darnell, R. B., RNA processing and its regulation: global insights into biological networks. *Nat. Rev. Genet.* **2010,** *11*, 75.

3. Gloss, B. S.; Dinger, M. E., Realizing the significance of noncoding functionality in clinical genomics. *Exp. Mol. Med.* **2018,** *50* (8), 97.

4. Tan, Z.-J.; Chen, S.-J., Nucleic Acid Helix Stability: Effects of Salt Concentration, Cation Valence and Size, and Chain Length. *Biophys. J.* **2006,** *90* (4), 1175-1190.

5. García-Ramos, J. C., et al., Metal-Based Drug-DNA Interactions. *J. Mex. Chem. Soc.* **2013,** *57*, 245-259.

6. Ganser, L. R., et al., The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.* **2019,** *20* (8), 474-489.

7. Wan, Y., et al., Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* **2014,** *505*, 706.

8. Sanchez de Groot, N., et al., RNA structure drives interaction with proteins. *Nat. Commun.* **2019,** *10* (1), 3246.

9. Wu, Y., et al., RNAex: an RNA secondary structure prediction server enhanced by high-throughput structure-probing data. *Nucleic Acids Res.* **2016,** *44* (W1), W294-W301.

10. Eddy, S. R., Non–coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* **2001,** *2* (12), 919-929.

11. Proudfoot, N. J., et al., Integrating mRNA Processing with Transcription. *Cell* **2002,** *108* (4), 501-512.

12. Lee, Y.; Rio, D. C., Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu. Rev. Biochem.* **2015,** *84* (1), 291-323.

13. Baralle, F. E.; Giudice, J., Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* **2017,** *18*, 437.

14. Treiber, T., et al., Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nat. Rev. Mol. Cell Biol.* **2019,** *20* (1), 5-20.

15. Sevignani, C., et al., Mammalian microRNAs: a small world for fine-tuning gene expression. *Mamm. Genome* **2006,** *17* (3), 189-202.

16. Paul, P., et al., Interplay between miRNAs and human diseases. *J. Cell. Physiol.* **2018,** *233* (3), 2007-2018.

17. Zhang, B., et al., microRNAs as oncogenes and tumor suppressors. *Dev. Biol.* **2007,** *302* (1), 1-12.

18. Iorio, M. V., et al., Breast cancer and microRNAs: therapeutic impact. *The Breast* **2011,** *20*, S63-S70.

19. Ruan, K., et al., MicroRNAs: Novel regulators in the hallmarks of human cancer. *Cancer Lett.* **2009,** *285* (2), 116-126.

20. Cooper, T. A., et al., RNA and Disease. *Cell* **2009,** *136* (4), 777-793.

21. Moszyńska, A., et al., SNPs in microRNA target sites and their potential role in human disease. *Open Biol. 7* (4), 170019.

22. Yue, M., et al., MSDD: a manually curated database of experimentally supported associations among miRNAs, SNPs and human diseases. *Nucleic Acids Res.* **2017,** *46* (D1), D181-D185.

23. Emilsson, V., et al., Genetics of gene expression and its effect on disease. *Nature* **2008,** *452* (7186), 423-428.

24. Kumar, A., et al., The Use of RNA Sequencing to Identify Disease-Specific Gene Expression Signatures and Critical Regulatory Networks Across Hematologic Malignancies. *Blood* **2014,** *124* (21), 2203.

25. Liu, C., et al., From Gene Expression To Disease Association. *Eur. Neuropsychopharmacol.* **2017,** *27*, S416.

26. Kim, S., et al., microRNA-155 positively regulates glucose metabolism via PIK3R1-FOXO3a-cMYC axis in breast cancer. *Oncogene* **2018,** *37* (22), 2982-2991.

27. Mattiske, S., et al., The Oncogenic Role of miR-155 in Breast Cancer. *Cancer Epidemiol., Biomarkers Prev.* **2012,** *21* (8), 1236.

28. Bustin, S. A., et al., The MIQE Guidelines: Information for Publication of Quantitative Real -Time PCR Experiments. *Clin. Chem.* **2009,** *55* (4), 611.

29. Nolan, T., et al., Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **2006,** *1* (3), 1559-1582.

30. Dheda, K., et al., The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. *Anal. Biochem.* **2005,** *344* (1), 141-143.

31. Schwarzenbach, H., et al., Data Normalization Strategies for MicroRNA Quantification. *Clin. Chem.* **2015,** *61* (11), 1333.

32. Honda, S.; Kirino, Y., Dumbbell-PCR: a method to quantify specific small RNA variants with a single nucleotide resolution at terminal sequences. *Nucleic Acids Res.* **2015,** *43* (12), e77-e77.

33. Pfaffl, M. W., *A-Z of Quantitative PCR*. International University Line: La Jolla, CA, 2004.

34. Forootan, A., et al., Methods to determine limit of detection and limit of quantification in quantitative real-time PCR (qPCR). *Biomol. Detect. Quantif.* **2017,** *12*, 1-6.

35. Taylor, S. C., et al., The Ultimate qPCR Experiment: Producing Publication Quality, Reproducible Data the First Time. *Trends Biotechnol.* **2019,** *37* (7), 761-774.

36. Streit, S., et al., Northern blot analysis for detection and quantification of RNA in pancreatic cancer cells and tissues. *Nat. Protoc.* **2008,** *4*, 37.

37. Pall, G. S.; Hamilton, A. J., Improved northern blot method for enhanced detection of small RNA. *Nat. Protoc.* **2008,** *3* (6), 1077-1084.

38. Choi, Y. S., et al., Removing bias against short sequences enables northern blotting to better complement RNA-seq for the study of small RNAs. *Nucleic Acids Res.* **2017,** *45* (10), e87-e87.

39. Zajc, K., et al., Nonradioactive northern blotting for the determination of acetylcholinesterase mRNA. Comparison to the radioactive technique. *Pfluegers Arch.* **2000,** *439* (7), R66-R67.

40. He, S. L.; Green, R., Chapter Three - Northern Blotting. In *Methods in Enzymology*, Lorsch, J., Ed. Academic Press: 2013; Vol. 530, pp 75-87.

41. Gall, J. G.; Pardue, M. L., FORMATION AND DETECTION OF RNA-DNA HYBRID MOLECULES IN CYTOLOGICAL PREPARATIONS. *Proc. Natl. Acad. Sci.* **1969,** *63* (2), 378.

42. Huber, D., et al., Fluorescence in situ hybridization (FISH): History, limitations and what to expect from micro-scale FISH? *Micro and Nano Engineering* **2018,** *1*, 15-24.

43. Jensen, E., Technical Review: In Situ Hybridization. *The Anatomical Record* **2014,** *297* (8), 1349-1353.

44. Choi, H. M. T., et al., Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat. Biotechnol.* **2010,** *28*, 1208.

45. Choi, H. M. T., et al., Next-Generation in Situ Hybridization Chain Reaction: Higher Gain, Lower Cost, Greater Durability. *ACS Nano* **2014,** *8* (5), 4284-4294.

46. Huang, K., et al., Quantitative, super-resolution localization of small RNAs with sRNA-PAINT. *bioRxiv* **2019**, 716696.

47. Urbanek, O. M., et al., Small RNA Detection by in Situ Hybridization Methods. *Int. J. Mol. Sci.* **2015,** *16* (6).

48. Maher, C. A., et al., Transcriptome sequencing to detect gene fusions in cancer. *Nature* **2009,** *458* (7234), 97-101.

49. Monroy-Contreras, R.; Vaca, L., Molecular Beacons: Powerful Tools for Imaging RNA in Living Cells. *J. Nucleic Acids* **2011,** *2011*, 15.

50. Pun, S. H.; Hoffman, A. S., B.8 - Nucleic Acid Delivery. In *Biomaterials Science (Third Edition)*, Ratner, B. D.; Hoffman, A. S.; Schoen, F. J.; Lemons, J. E., Eds. Academic Press: 2013; pp 1047-1054.

51. Chesnoy, S.; Huang, L., Structure and Function of Lipid-DNA Complexes for Gene Delivery. *Annu. Rev. Biophys. Biomol. Struct.* **2000,** *29* (1), 27-47.

52. Felgner, P. L., et al., Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci.* **1987,** *84* (21), 7413.

53. Hemphill, J.; Deiters, A., DNA Computation in Mammalian Cells: MicroRNA Logic Operations. *J. Am. Chem. Soc.* **2013,** *135* (28), 10512-10518.

54. Wu, C., et al., A Nonenzymatic Hairpin DNA Cascade Reaction Provides High Signal Gain of mRNA Imaging inside Live Cells. *J. Am. Chem. Soc.* **2015,** *137* (15), 4900-4903.

55. Cheglakov, Z., et al., Live Cell MicroRNA Imaging Using Cascade Hybridization Reaction. *J. Am. Chem. Soc.* **2015,** *137* (19), 6116-6119.

56. Barber, M. A., A Technic for the Inoculation of Bacteria and Other Substances into Living Cells. *J. Infect. Dis.* **1911,** *8* (3), 348-360.

57. Gurdon, J. B., et al., Use of Frog Eggs and Oocytes for the Study of Messenger RNA and its Translation in Living Cells. *Nature* **1971,** *233* (5316), 177-182.

58. Jinturkar, K. A., et al., 3 - Gene Delivery Using Physical Methods. In *Challenges in Delivery of Therapeutic Genomics and Proteomics*, Misra, A., Ed. Elsevier: London, 2011; pp 83-126.

59. Chen, A. K., et al., Ratiometric bimolecular beacons for the sensitive detection of RNA in single living cells. *Nucleic Acids Res.* **2010,** *38* (14), e148-e148.

60. Chen, A. K., et al., Avoiding false-positive signals with nuclease-vulnerable molecular beacons in single living cells. *Nucleic Acids Res.* **2007,** *35* (16), e105-e105.

61. Green, M.; Loewenstein, P. M., Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* **1988,** *55* (6), 1179-1188.

62. Frankel, A. D.; Pabo, C. O., Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **1988,** *55* (6), 1189-1193.

63. Jeong, C., et al., A branched TAT cell-penetrating peptide as a novel delivery carrier for the efficient gene transfection. *Biomater. Res.* **2016,** *20* (1), 28.

64. Kim, H. H., et al., Basic peptide system for efficient delivery of foreign genes. *Biochim. Biophys. Acta, Mol. Cell Res.* **2003,** *1640* (2), 129-136.

65. Abes, R., et al., Cell-penetrating-peptide-based delivery of oligonucleotides: an overview. *Biochem. Soc. Trans.* **2007,** *35* (4), 775.

66. Nitin, N., et al., Peptide-linked molecular beacons for efficient delivery and rapid mRNA detection in living cells. *Nucleic Acids Res.* **2004,** *32* (6), e58-e58.

67. Chen, A. K., et al., Efficient cytosolic delivery of molecular beacon conjugates and flow cytometric analysis of target RNA. *Nucleic Acids Res.* **2008,** *36* (12), e69-e69.

68. Chen, A. K., et al., Sub-cellular trafficking and functionality of 2′- O -methyl and 2′- O -methyl-phosphorothioate molecular beacons. *Nucleic Acids Res.* **2009,** *37* (22), e149-e149.

69. Molenaar, C., et al., Linear 2' O-Methyl RNA probes for the visualization of RNA in living cells. *Nucleic Acids Res.* **2001,** *29* (17), e89.

70. Seferos, D. S., et al., Nano-Flares: Probes for Transfection and mRNA Detection in Living Cells. *J. Am. Chem. Soc.* **2007,** *129* (50), 15477-15479.

71. Peng, H., et al., A microRNA-initiated DNAzyme motor operating in living cells. *Nat. Commun.* **2017,** *8* (1), 14378.

72. Wu, Y., et al., DNA aptamer–micelle as an efficient detection/delivery vehicle toward cancer cells. *Proc. Natl. Acad. Sci.* **2010,** *107* (1), 5.

73. Qiu, L., et al., A Targeted, Self-Delivered, and Photocontrolled Molecular Beacon for mRNA Detection in Living Cells. *J. Am. Chem. Soc.* **2013,** *135* (35), 12952-12955.

74. Arvey, A., et al., Minimizing off-target signals in RNA fluorescent in situ hybridization. *Nucleic Acids Res.* **2010,** *38* (10), e115-e115.

75. Li, J. J., et al., Molecular Beacons: A Novel Approach to Detect Protein – DNA Interactions. *Angew. Chem., Int. Ed.* **2000,** *39* (6), 1049-1052.

76. Wang, L., et al., Locked Nucleic Acid Molecular Beacons. *J. Am. Chem. Soc.* **2005,** *127* (45), 15664-15665.

77. Mao, C., et al., A nanomechanical device based on the B–Z transition of DNA. *Nature* **1999,** *397* (6715), 144-146.

78. Yurke, B., et al., A DNA-fuelled molecular machine made of DNA. *Nature* **2000,** *406* (6796), 605-608.

79. Lee, C. S., et al., A physical study by electron microscopy of the terminally repetitious, circularly permuted DNA from the coliphage particles of Escherichia coli 15. *J. Mol. Biol.* **1970,** *48* (1), 1-22.

80. Seelig, G., et al., Enzyme-Free Nucleic Acid Logic Circuits. *Science* **2006,** *314* (5805), 1585.

81. Qian, L.; Winfree, E., Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **2011,** *332* (6034), 1196.

82. Phillips, A.; Cardelli, L., A programming language for composable DNA circuits. *J. R. Soc., Interface* **2009,** *6* (suppl_4), S419-S436.

83. Lakin, M. R., et al., Visual DSD: a design and analysis tool for DNA strand displacement systems. *Bioinformatics* **2011,** *27* (22), 3211-3213.

84. Zhang, D. Y.; Winfree, E., Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.* **2009,** *131* (47), 17303-17314.

85. Srinivas, N., et al., Enzyme-free nucleic acid dynamical systems. *Science* **2017,** *358* (6369), eaal2052.

86. Kotani, S.; Hughes, W. L., Multi-Arm Junctions for Dynamic DNA Nanotechnology. *J. Am. Chem. Soc.* **2017,** *139* (18), 6363-6368.

87. Groves, B., et al., Computing in mammalian cells with nucleic acid strand exchange. *Nat. Nanotechnol.* **2015,** *11*, 287-295.

88. Khvorova, A.; Watts, J. K., The chemical evolution of oligonucleotide therapies of clinical utility. *Nat. Biotechnol.* **2017,** *35*, 238-248.

89. Robbins, M., et al., 2′-O-methyl-modified RNAs Act as TLR7 Antagonists. *Mol. Ther.* **2007,** *15* (9), 1663-1669.

90. Lennox, K. A., et al., Characterization of Modified Antisense Oligonucleotides in Xenopus laevis Embryos. *Oligonucleotides* **2006,** *16*, 26-42.

91. Shen, W., et al., Acute hepatotoxicity of 2′ fluoro-modified 5–10–5 gapmer phosphorothioate oligonucleotides in mice correlates with intracellular protein binding and the loss of DBHS proteins. *Nucleic Acids Res.* **2018,** *46* (5), 2204-2217.

92. Seth, P. P., et al., Structure Activity Relationships of alpha-L-LNA Modified Phosphorothioate Gapmer Antisense Oligonucleotides in Animals. *Mol. Ther. - Nucleic Acids* **2012,** *1*, e47.

93. Afonin, K. A., et al., Activation of different split functionalities on re-association of RNA–DNA hybrids. *Nat. Nanotechnol.* **2013,** *8*, 296.

94. Wolfrum, C., et al., Mechanisms and optimization of in vivo delivery of lipophilic siRNAs. *Nat. Biotechnol.* **2007,** *25*, 1149.

95. Soutschek, J., et al., Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* **2004,** *432* (7014), 173-178.

96. Nair, J. K., et al., Multivalent N-Acetylgalactosamine-Conjugated siRNA Localizes in Hepatocytes and Elicits Robust RNAi-Mediated Gene Silencing. *J. Am. Chem. Soc.* **2014,** *136* (49), 16958-16961.

97. Letsinger, R. L.; Mahadevan, V., Oligonucleotide Synthesis on a Polymer Support1,2. *J. Am. Chem. Soc.* **1965,** *87* (15), 3526-3527.

98. Letsinger, R. L., et al., Nucleotide chemistry. XX. Phosphite coupling procedure for generating internucleotide links. *J. Am. Chem. Soc.* **1975,** *97* (11), 3278-3279.

99. Letsinger, R. L.; Lunsford, W. B., Synthesis of thymidine oligonucleotides by phosphite triester intermediates. *J. Am. Chem. Soc.* **1976,** *98* (12), 3655-3661.

100. Beaucage, S. L.; Caruthers, M. H., Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **1981,** *22* (20), 1859-1862.

101. Matteucci, M. D.; Caruthers, M. H., Synthesis of deoxyoligonucleotides on a polymer support. *J. Am. Chem. Soc.* **1981,** *103* (11), 3185-3191.

102. Ashley, G. W., Modeling, Synthesis, and Hybridization Properties of (L)-Ribonucleic Acid. *J. Am. Chem. Soc.* **1992,** *114* (25), 9731-9736.

103. Hoehlig, K., et al., Stereospecificity of oligonucleotide interactions revisited: no evidence for heterochiral hybridization and ribozyme/DNAzyme activity. *PLoS One* **2015,** *10* (2), e0115328.

104. Urata, H., et al., Synthesis and properties of mirror-image DNA. *Nucleic Acids Res.* **1992,** *20* (13), 3325-3332.

105. Hauser, N. C., et al., Utilising the left-helical conformation of L-DNA for analysing different marker types on a single universal microarray platform. *Nucleic acids Res.* **2006,** *34* (18), 5101-5111.

106. Garbesi, A., et al., L-DNA as potential antimessenger oligonucleotides: a reassessment. *Nucleic Acids Res.* **1993,** *21* (18), 4159-4165.

107. Egholm, M., et al., Peptide nucleic acids (PNA). Oligonucleotide analogs with an achiral peptide backbone. *J. Am. Chem. Soc.* **1992,** *114* (5), 1895-1897.

108. Egholm, M., et al., Recognition of guanine and adenine in DNA by cytosine and thymine containing peptide nucleic acids (PNA). *J. Am. Chem. Soc.* **1992,** *114* (24), 9677-9678.

109. Jensen, K. K., et al., Kinetics for Hybridization of Peptide Nucleic Acids (PNA) with DNA and RNA Studied with the BIAcore Technique. *Biochemistry* **1997,** *36* (16), 5072-5077.

110. Egholm, M., et al., PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules. *Nature* **1993,** *365* (6446), 566-568.

111. Pellestor, F.; Paulasova, P., The peptide nucleic acids (PNAs), powerful tools for molecular genetics and cytogenetics. *Eur. J. Hum. Genet.* **2004,** *12* (9), 694-700.

112. Wittung, P., et al., DNA-like double helix formed by peptide nucleic acid. *Nature* **1994,** *368* (6471), 561-563.

113. Brown, S. C., et al., NMR solution structure of a peptide nucleic acid complexed with RNA. *Science* **1994,** *265* (5173), 777.

114. Gildea, B. D., et al., PNA solubility enhancers. *Tetrahedron Lett.* **1998,** *39* (40), 7255-7258.

# CHAPTER II

## DEVELOPMENT OF HETEROCHIRAL STRAND DISPLACEMENT REACTIONS*

Dynamic DNA nanodevices, such as those described in section 1.5.1, almost invariably represent homochiral systems composed exclusively of D-DNA, the naturally occurring stereoisomer. This is despite the obvious attractiveness of L-DNA as a biopolymer for the design of biologically stable strand displacement cascades and circuits. The key challenge associated with integrating both enantiomers of DNA into a single device is their inability to form contiguous WC base pairs with each other[1]. While this property alone is potentially beneficial due to decreased non-specific interactions with natural oligonucleotides (ONs), it precludes the sequence-specific transfer of information between the two enantiomers of DNA. In this fashion, L-DNA logic operations cannot be linked to the presence of a native D-ON target (generally an RNA target, in the context of relevant single-stranded nucleic acid species in live cells).

Here, we reasoned that this limitation could be overcome by employing an achiral nucleic acid analogue as a sequence-specific mediator between the two orthogonal enantiomers of DNA. For this role, we turned our attention to PNA[2-4]. As a reference, the monomeric subunits of D-DNA, L-DNA and PNA are shown in Figure II-1. Since PNA obeys canonical WC base pairing rules, but has no inherent chirality, it can hybridize to either D or L-ONs with equal affinity[5-6]. Based on this property, we conceived of a novel toehold-mediate strand-displacement reaction exploiting a PNA/DNA heteroduplex (Figure II-2) in order to interface the two enantiomers of DNA[7]. Although the

Adapted with permission from Heterochiral DNA Strand-Displacement Circuits, by Kabza *et al*, 2017, *J. Am. Chem. Soc.*, 139 (49), 17715-17718.Copyright 2017 American Chemical Society.*
Adapted with permission from Heterochiral DNA Strand-Displacement Based on Chimeric D/L-Oligonucleotides, by Young, B and Sczepanski, J, 2019, *ACS Synth. Biol.*, doi.org/10.1021/acssynbio.9b00335.Copyright 2019 American Chemical Society.*

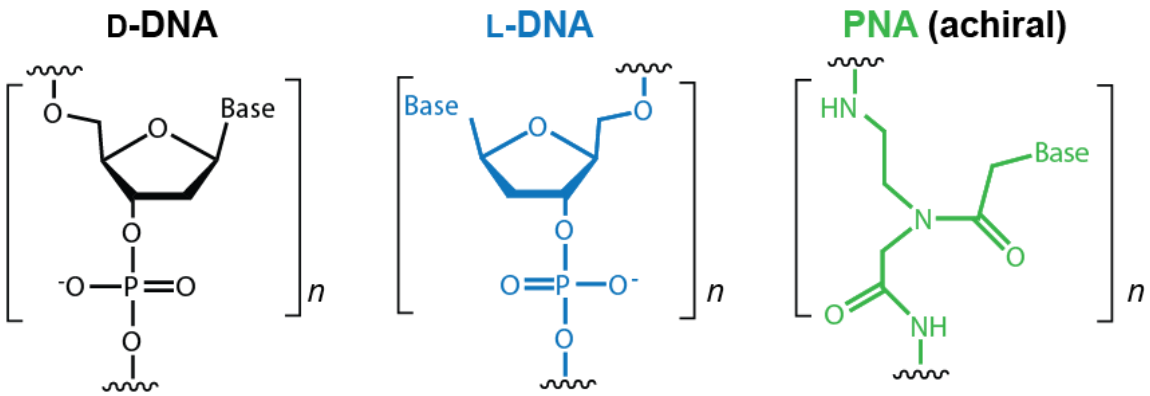



**Figure II-1. D-DNA, L-DNA and PNA monomers.** The three types of monomers utilized in the design of heterochiral strand displacement systems. Sequences described in the following chapters will follow this color scheme. D-DNA will be black, L-DNA will be blue and PNA will be green.

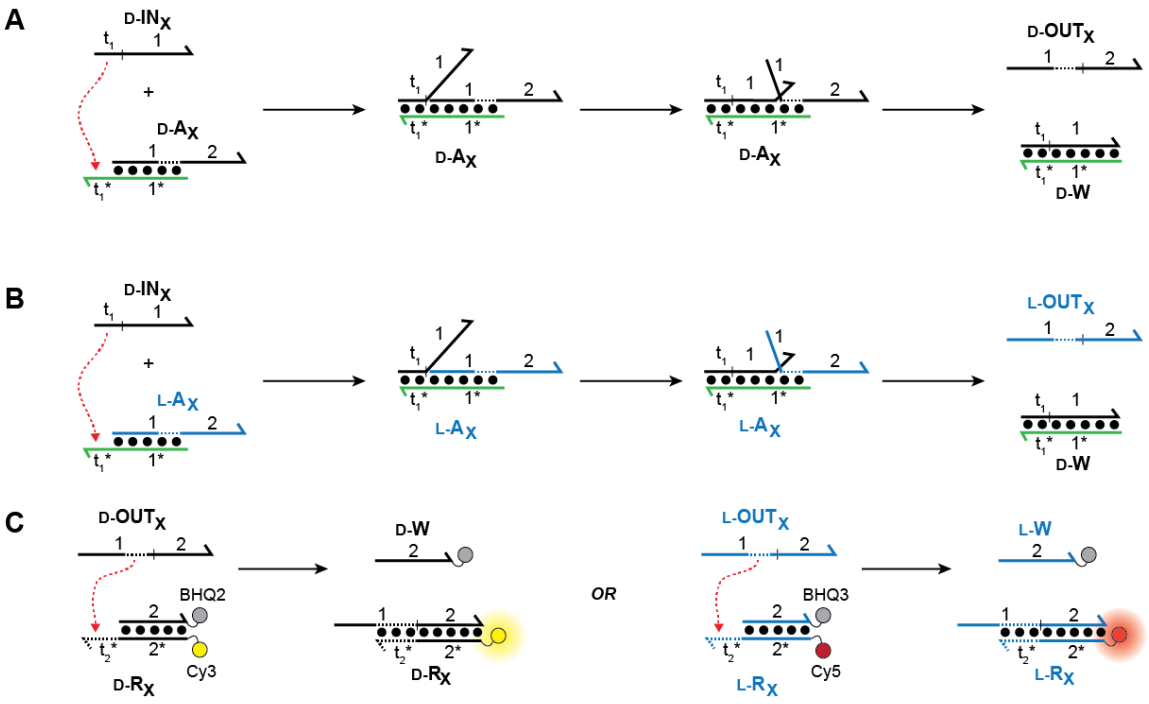



**Figure II-2. Strand displacement and domain notation.** A generalizable scheme for all strand displacements discussed in this work. Domains represent unique DNA sequences. Each $t_x$ represent a new toehold, each number represent a new domain, and complements are indicated by an asterisk (*). The 3′ end of each strand is denoted by a half-arrow. (A) Homochiral strand displacement. The input strand is the same chirality as the incumbent strand (OUT) hybridized to the PNA. (B) Heterochiral strand displacement. The input strand is the opposite chirality as the incumbent strand (OUT) hybridized to the PNA. (C) General reporter design and activation. Each OUT strand contains a new toehold that is sequestered within the duplex region when bound to the PNA (indicated in all figures as a dashed line segment). Once released from the PNA by strand displacement, this new toehold can bind to reporter duplexes and activate them by strand displacement. Inputs will be denoted as D or L-IN$_X$, inversion gates as D or L-A$_X$, and reporters as D or L-R$_X$. In each case, X will indicate the sequences used as outlined in appendix A. Waste strands or duplexes are assumed to not participate in the reaction, and are denoted by the letter W.

thermodynamic and kinetic parameters of PNA hybridizing to DNA are much different than natural DNA/DNA hybridization, these strand displacement reactions were loosely based on traditional DNA strand displacement (Figure II-2 A, B and C). For consistency, inputs for this style of strand displacement reaction will be referred to as either "homochiral" or "heterochiral" for the remainder of this dissertation. A homochiral input represents a strand displacement reaction where the invading strand is the same chirality as the incumbent strand hybridized to the PNA (Figure II-2 A). A heterochiral input will have the opposite chirality as the incumbent strand hybridized to the PNA (Figure II-2 B). These strand displacement reactions were monitored using a downstream reporter (Figure II-2 C). In the sections below we will provide an answer for a key question in this work: will a PNA toehold be a satisfactory substrate for nucleation and branch migration with a heterochiral DNA or RNA input.

## 2.1 Results

### *2.1.1 Design of heterochiral strand displacement systems based on PNA/DNA heteroduplexes*

To test the above hypothesis, a PNA was designed to be complementary to mature miR-155 (hereafter referred to as $PNA_{155}$), and both D and L-DNA strands that were the same sequence as miR-155 were synthesized by standard phosphoramidite chemistry. MiR-155 represents an attractive target for preliminary studies, since it is strongly overexpressed in some breast cancer cell lines[8-9]. DNA inputs were tested initially because they are easier to synthesize and handle than their RNA counterparts. D or L- DNA output

strands were designed to partially hybridize to $PNA_{155}$ (forming complex D or L-$A_{155}$, Figure II-2 A and B), and these output strands contain excess single stranded DNA to allow for toehold nucleation and branch migration through a downstream reporter (D or L-$R_{155}$, Figure II-2 C) after displacement from the PNA. The single-stranded domain of the $OUT_{155}$ strand and the reporter were computationally designed to have a low risk of secondary structure formation, and proper hybridization was validated by the online DNA structure prediction tool NUPACK[10]. Briefly, NUPACK allows users to submit code defining DNA or RNA strand length, sequence constraints and random bases that can be computationally defined (i.e. CTAATCGTGATAGGNNNNNNNNNNNNNN, where the given sequence will be complementary to the PNA and N can be either A, T, C or G). Additionally, the code allows the user to define the desired secondary structure, and NUPACK will predict bases within the random region that are expected to accommodate the chosen secondary structure. Once an output sequence is chosen from the list provided, the reporter sequences are designed to be complementary to the OUT strand as needed. An example version of this code is provided in Appendix B.

This system represents a "chirality OR" gate, with complex $A_{155}$ (D or L) acting as a novel translation gate (discussed in section 1.5.1) that can be invaded by an input of either chirality (D "OR" L). We've termed this type of translation gate an inversion gate, as it can invert the stereochemistry of the input strand from D to L or vice versa. Utilizing complexes L-$A_{155}$ and L-$R_{155}$, the performance of this gate was validated by adding either the D-DNA input, the L-DNA input, or both inputs (again, these sequences are identical to mature miR-155) and monitored the generation of fluorescent signal (Figure II-3 B).

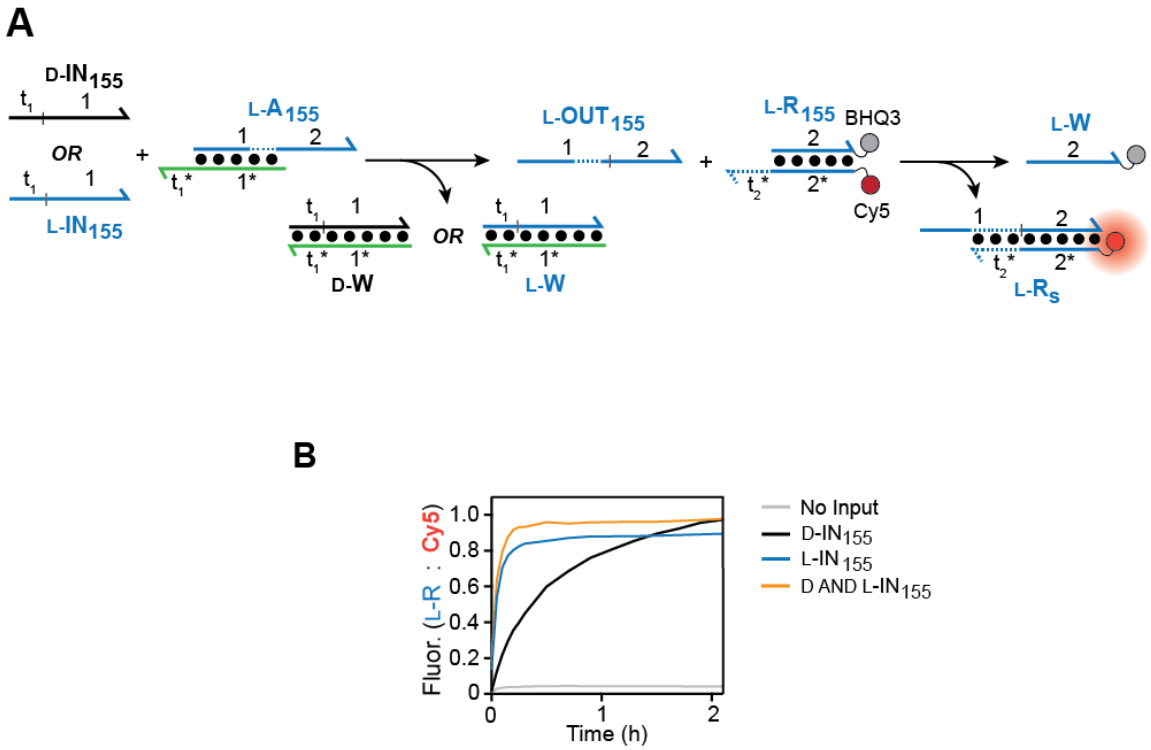



**Figure II-3. A chirality "OR" gate with L-A155.** (A) Schematic of homochiral and heterochiral strand displacements with D or L-DNA inputs the same sequence as miR-155. (B) Reaction progress was monitored by the displacement of the quencher strand from L-$R_{155}$ over time. Fluorescence was normalized to the maximum achievable signal of the system, and each trace represents identical reaction conditions with the indicated inputs. Adapted with permission from Ref. 7. Copyright 2017 American Chemical Society.

Regardless of the input chirality, this inversion gate design generated maximum signal after 2 hours. All fluorescence traces were normalized to a positive control representing the maximum achievable signal for the system, minor differences in the observed endpoints of these reactions were likely due to pipetting error involving the $A_{155}$ duplex as this is the limiting reagent in these reactions. Interestingly, the chirality OR gate functioned significantly slower when activated with a heterochiral input (Figure II-3 B, D-IN). Since the thermodynamic endpoint of each reaction is the same regardless of input chirality (i.e. a fully complementary PNA/DNA waste duplex is formed, and the reporter is fully displaced), there appears to be some kinetic penalty incurred specifically when the heterochiral input was used. We hypothesized that this might arise from alterations in toehold nucleation between the homochiral or heterochiral systems, or a decreased rate of branch migration in the heterochiral strand displacement pathway. Considering branch migration is generally considered the 'slow step' with respect to the kinetics of strand displacement reactions[11], it seems reasonable that a kinetic penalty at this step would more readily manifest as a slower overall reaction. Additionally, the single stranded PNA toeholds in this system are expected to be achiral and as such the hybridization kinetics of either the D or L-input to the toehold should be equivalent. However, there is some literature suggesting that single unpaired bases at the end of DNA duplexes still participate in base stacking with the duplex[12], and it is interesting to consider that the chirality of the duplex may propagate in part into the single strand toehold region[13-15]. These observations will be discussed in further detail in Chapter IV.

Having demonstrated the capability of heterochiral strand displacement, and since the PNA used in these experiments had been designed as the complement to miR-155, we wanted to evaluate the performance of this system with an RNA input in the presence or absence of total RNA extract from HeLa cells. In this way we expected to show the specificity of our system for the sequence-specific target in the presence of a large excess of non-specific RNA (Figure II-4). As expected, a D-RNA input was able to undergo heterochiral strand displacement and release an L-DNA output strand, providing the first potential strategy for autonomously interfacing endogenous nucleic acids with nanodevices composed primarily of bio-orthogonal L-DNA. Additionally, a heterochiral RNA input reacted faster with L-$A_{155}$ than the corresponding heterochiral DNA input of the same sequence. This was not exactly surprising, as RNA has been reported to hybridize to PNA with faster $k_{on}$ and slower $k_{off}$ rates than the corresponding DNA to PNA hybridization[16], but it has interesting implications for the use of these systems for the detection of RNA targets in live cells.

We further characterized the kinetics of these strand displacement reaction with each of the tested inputs: L-IN, D-IN and D-$IN_{RNA}$. The kinetics of DNA strand displacement reactions are approximated as 2$^{nd}$-order with respect to the input and the inversion gate[17-18]. This assumption is made because the rate of the reporter reaction with the output strand is experimentally designed to be faster than the other steps in the system, often by using the reporter duplex in excess compared to other reaction components or increasing the length of its toehold domain. In all reactions described in this chapter the reporter was used in 3-fold excess with respect to the inversion gate, and experimentally

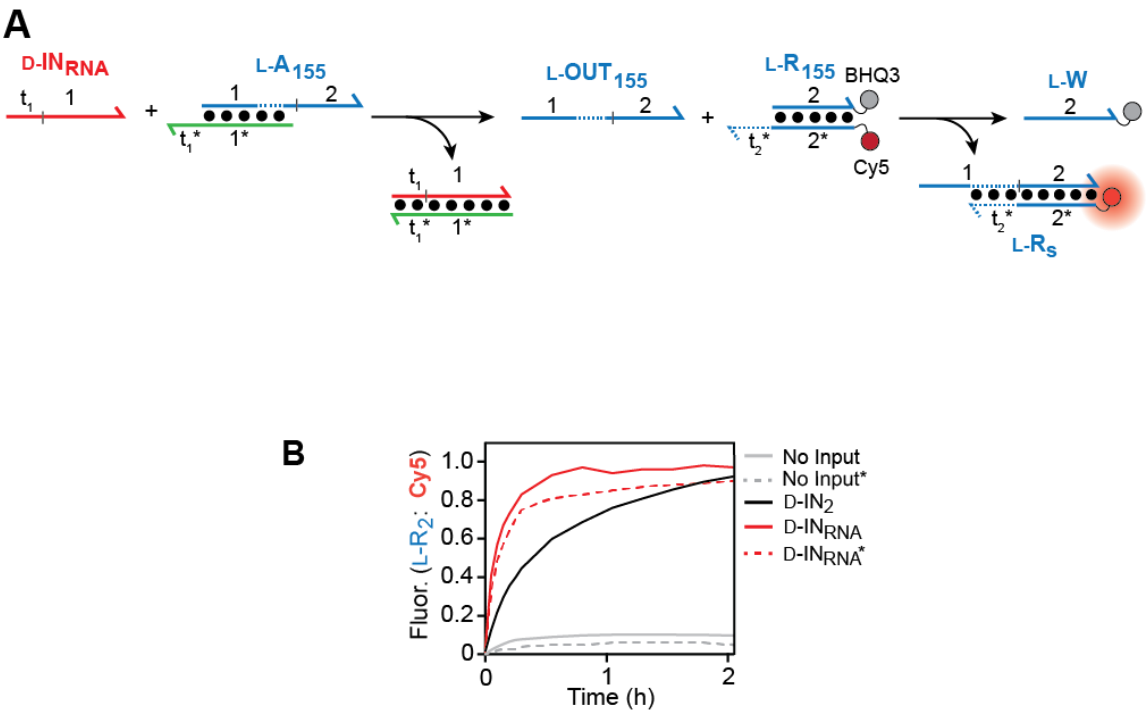



**Figure II-4. Heterochiral strand displacement with miR-155.** (A) Scheme for heterochiral strand displacement using the miR-155 input. (B) Traces represent the specific activation of L-A$_{155}$ by miR-155 in the absence (solid line) or presence (dashed line) of 0.1 mg/mL HeLa cell nuclear RNA extract. Adapted with permission from Ref. 7. Copyright 2017 American Chemical Society.

we have observed that strand displacement reactions proceeding through a PNA toehold are much slower than the corresponding all DNA strand displacements (like the reporter reaction). Normalized reaction traces were fit to rate equations as previously described[17] and rate constants were calculated to be $2.14x10^4$ $M^{-1}$ $s^{-1}$, $9.56x10^2$ $M^{-1}$ $s^{-1}$, and $5.13x10^3$ $M^{-1}$ $s^{-1}$ for L-IN, D-IN and D-$IN_{RNA}$ respectively. Even strand displacement by the homochiral input was substantially slower than reported rates for all-DNA strand displacement reactions (~$10^4$ vs. ~$10^6$)[18-19], despite the fact that hybridization of DNA onto a PNA toehold should be faster than DNA onto a DNA toehold. This can be rationalized if we consider the more thermodynamically stable PNA/DNA branch migration domain as a sum of individual steps (i.e. each incumbent base pair that must be broken during strand invasion) each imposing a higher free energy barrier than the corresponding DNA/DNA counterpart.

Having shown capability of building a heterochiral strand displacement system capable of detecting a heterochiral miR target, we followed the same principles to design two new systems based on PNAs complimentary to regions of miR-10b as well as the MnSOD mRNA (Figure II-5 A and B, respectively). Again, these targets represent RNAs often overexpressed in various types of cancer[20-21], and were chosen in order to roughly test what effects increasing the toehold strength would have on these reaction systems. These reactions were initially tested with DNA inputs to verify strand displacement and general performance. All three of these systems generate a PNA/DNA waste duplexes with similar $T_m$s (73.9, 75.3, and 76.5 for the miR-155, miR-10b, and MnSOD systems, respectively), in an attempt to isolate the toehold effect as much as possible. Interestingly,

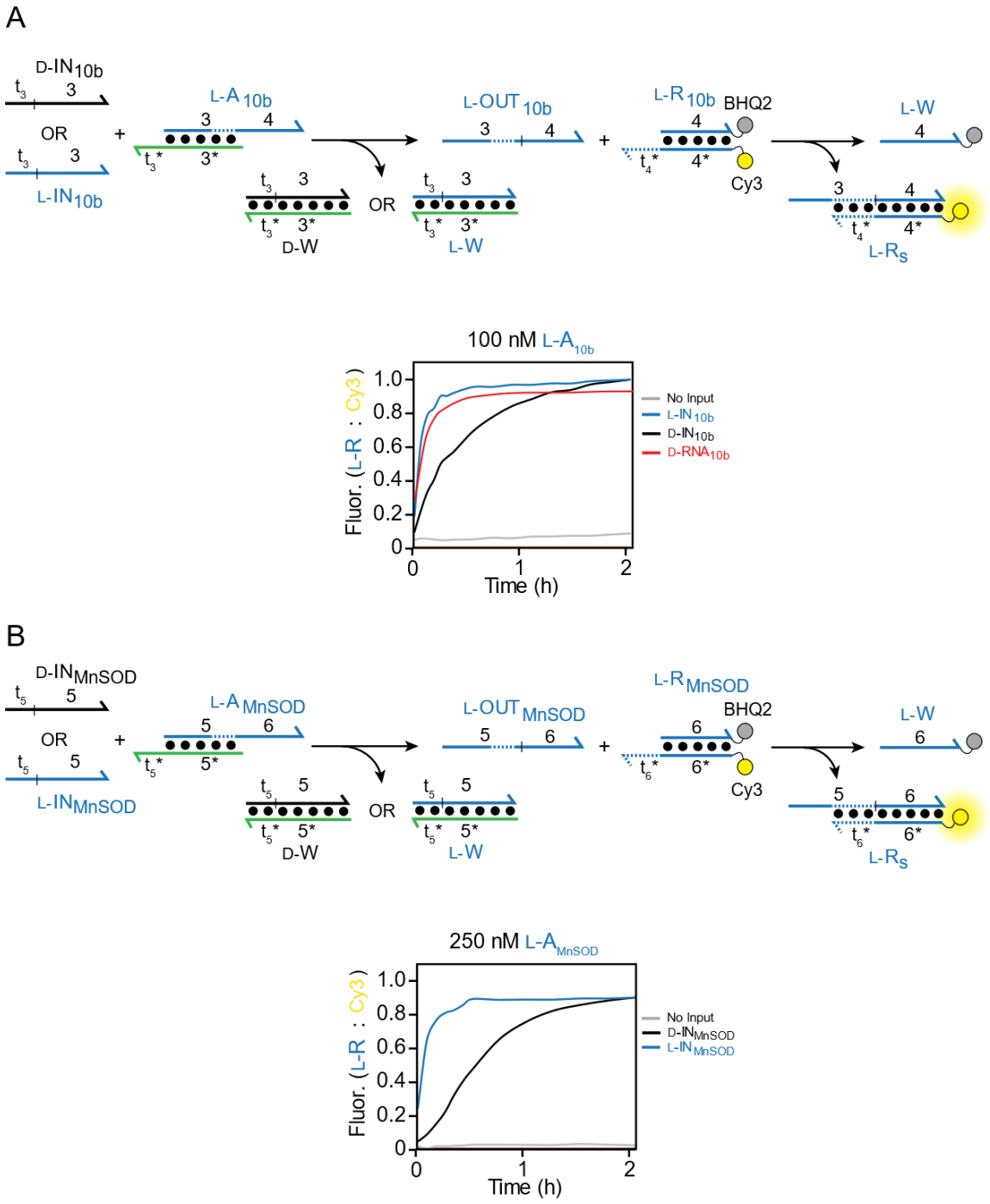



**Figure II-5. Heterochiral strand displacement systems for miR-10b and MnSOD mRNA.** (A) Scheme for the strand displacement of L-$A_{10b}$ by D-DNA, L-DNA or RNA inputs the same sequence as miR-10b. (B) Scheme for the strand displacement of L-$A_{MnSOD}$ by D or L-DNA inputs representing the target region of the MnSOD mRNA.

compared to the other two reactions discussed in this section the miR-10b inversion gate possesses a substantially faster reaction rate for both the homochiral and heterochiral inputs ($3.2x10^4$ $M^{-1}$ $s^{-1}$ and $4.4x10^3$ $M^{-1}$ $s^{-1}$, respectively). This is only a difference of ~7-fold in the rate constants of the homochiral and heterochiral reactions, as opposed to the >20-fold difference seen in the miR-155 system. This observation might arise from two potential sources, or likely a combination thereof: 1) the branch migration domain of the miR-10b inversion gate is the shortest of the ones tested (13 bases long), although this was roughly accounted for as previously described, and 2) the toehold of the miR-10b inversion gate contains 3 "strong" base pairs (G – C) while the other toeholds only contain one (miR-155) or two (MnSOD). This trend is consistent, with the MnSOD system being faster than the 155 system. The impact of strong base pairing within the toehold of all-DNA strand displacement systems has been partially described previously[18-19], with strand displacement rates varying over 2 orders of magnitude with toehold strength even at the same toehold length.

We next sought to expand the complexity of these heterochiral strand displacement systems by designing an all L-DNA AND gate that would be activated by two different inversion gates. The inversion gates designed for miR-155 and miR-10b were chosen as the basis of this reaction cascade (Figure II-6). In these experiments, successful displacement of L-$OUT_{10b}$ from duplex L-$A_{10b}$ displaces the blocking strand from the L-AND gate by branch migration (Figure II-6 A), revealing a new toehold. Then, displacement of L-$OUT_{155}$ from duplex L-$A_{155}$ leads to the displacement of the quencher-labeled strand from the L-AND gate (Figure II-6 B), generating fluorescence and

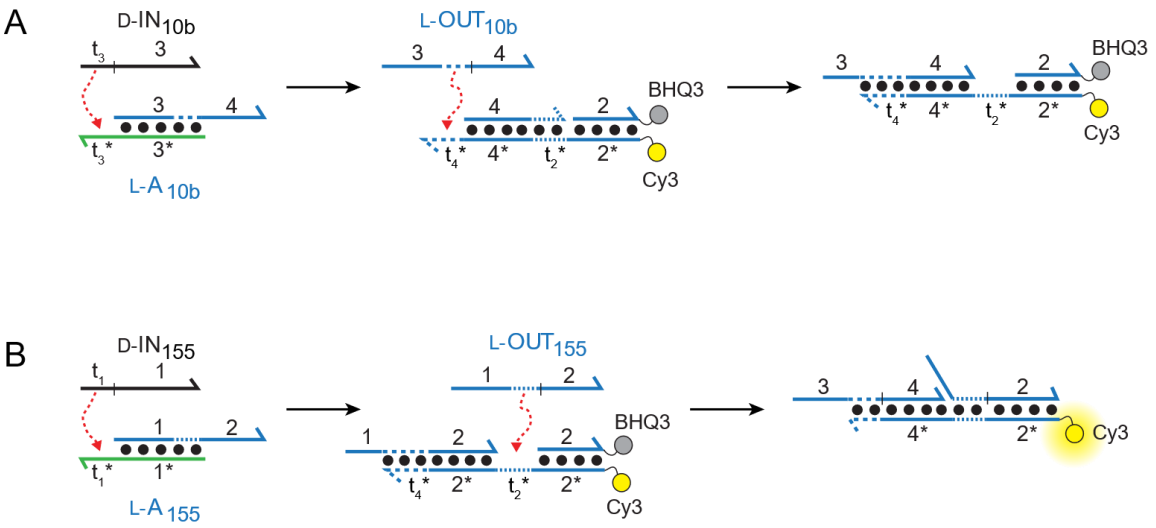



**Figure II-6. (D-IN$_{10b}$ "OR" L-IN$_{10b}$) "AND" (D-IN$_{155}$ "OR" L-IN$_{155}$).** (A) Displacement of L-OUT$_{10b}$ by D-IN$_{10b}$ (or L-IN$_{10b}$) reveals a new toehold (large dashes) within domain a that binds to the AND gate reporter and displaces domain b, revealing a new toehold on the AND gate (small dashed line). (B) Displacement of L-OUT$_{155}$ by D-IN$_{155}$ (or L-IN$_{155}$) reveals a new toehold within domain 1 that binds to the second toehold of the AND gate and displaces the quencher.

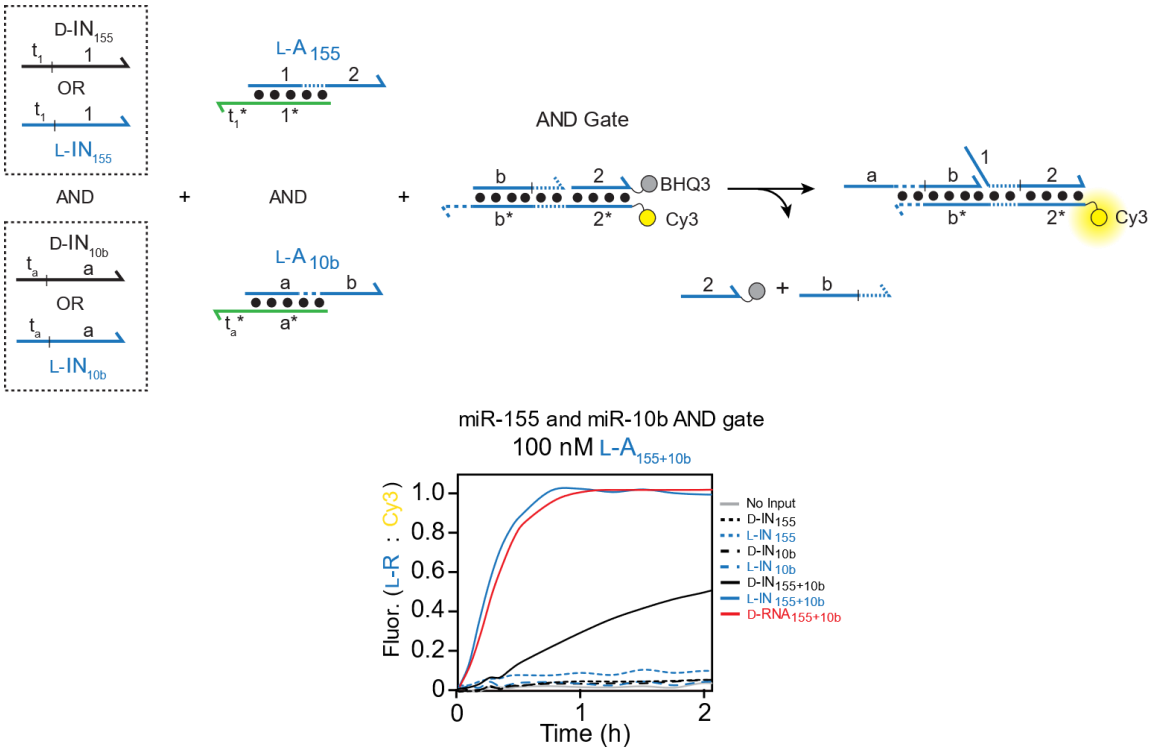



**Figure II-7. Activation of an L-AND gate by heterochiral strand displacement.** Scheme showing potential inputs for the AND gate. Activation was monitored by fluorescence; leak reactions are indicated by dashed lines and represent either L-IN$_{10b}$ or L-IN$_{155}$ alone (blue) or D-IN$_{10b}$ or D-IN$_{155}$ alone (black). Activation by D-RNA versions of miR-155 and miR-10b is shown in red.

indicating that all displacements have occurred. Again, the performance of the L-AND gate was validated with both the homochiral and heterochiral inputs for each inversion gate (Figure II-7) and monitored by spectrofluorimetry. While the apparent rates of both the heterochiral and homochiral displacements vary from system to system, their general trends remain the same, even in multicomponent system. All reactions are relatively fast when initiated with a homochiral input but are substantially slower when initiated by a heterochiral input.

### *2.1.2 Enabling DNA circuit thresholding with heterochiral strand displacement*

Reaction thresholding in DNA circuits is a method often used to prevent signal degradation and reduce leak reactions. This is most often accomplished by the addition of a thresholding duplex that competes with downstream components for DNA output strands[22]. These thresholding gates are generally at low concentration but are designed to be kinetically favored reactants, often by modulating the toehold length of these components. When a DNA circuit leaks, meaning the output strand is released in small quantities in the absence of the correct input, a threshold gate can be used to 'soak up' these strands and prevent spurious activation of the reporter complex. Because the rates of these strand displacement reactions can be modified by orders of magnitude based only on toehold length, small amounts of a threshold gate can effectively sequester the small amount of output released in these leak reactions. Since these thresholding reactions are modulated by the kinetics of strand displacement, we thought that our heterochiral strand displacement systems might be able to act like a threshold gate based on chirality, rather

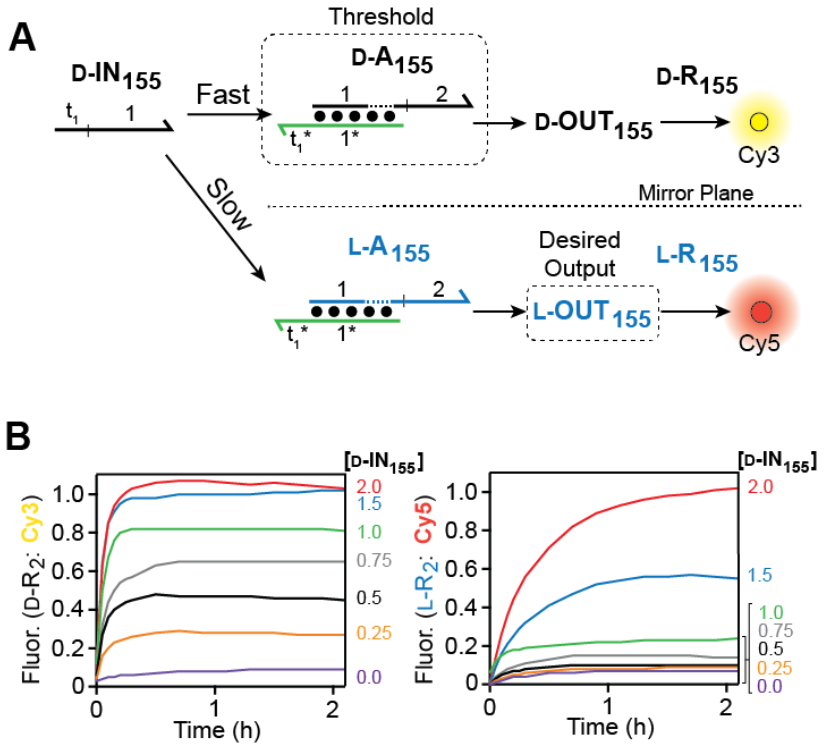



**Figure II-8. A "thresholding" gate based on the kinetics of heterochiral strand displacement.** For clarity, Cy3 and Cy5 spheres represent the observed fluorescence when either D or L-$R_{155}$ is activated, respectively. (A) Schematic illustrating the experiment design of the thresholding experiment. Based on the observed rates of strand displacement with the homochiral input being much faster than with the heterochiral input, we expected that if both D and L-$A_{155}$ were present in the same reaction mixture D-$IN_{155}$ would preferentially activate D-$R_{155}$ before L-R155. (B) Fluorescence of both reporters was monitored simultaneously with filters for Cy3 and Cy5 (D-$R_{155}$ and L-$R_{155}$, respectively). As expected, up to 1 equivalent of D-$IN_{155}$ activated D-$R_{155}$ almost exclusively. Excess D-$IN_{155}$ activates L-$R_{155}$ as expected.

than toehold length. As described in the previous section, the rate of strand displacement off a PNA/DNA heteroduplex is dramatically slower when activated by a heterochiral input. We hypothesized that a reaction mixture containing both D and L-$A_{155}$ duplexes, as well as both D and L-$R_{155}$, would preferentially react with a given input in a homochiral fashion. To test this, a series of strand displacement reactions containing both D and L-$A_{155}$, as well as D- and L-$R_{155}$, was prepared and these reactions were initiated with 0.25, 0.5, 0.75, 1.0, 1.5 or 2.0 equivalents (with respect to the D- and L-$A_{155}$ duplexes) of the D-input. In this experiment D-$R_{155}$ was labeled with Cy3 and L-$R_{155}$ was labeled with Cy5 so that reporter activation could be monitored simultaneously. Indeed, at input concentrations below that of the D-$A_{155}$ duplex ($\leq$ 1 equivalent) almost exclusive activation of D-$R_{155}$ was observed (Cy3 signal, Figure II-8). At input concentrations exceeding the D-$A_{155}$ duplex (> 1 equivalent), L-$R_{155}$ was activated as expected. The work here describes the first example of a DNA thresholding gate solely dependent on the chirality of the input.

### *2.1.3 Design of heterochiral strand displacement systems based on chimeric D/L-ONs*

PNA synthesis isn't compatible with standard DNA synthesis techniques, therefore individual PNA sequences and any modifications thereof need to be synthesized independently or ordered from commercial sources. We realized this process is expensive, time consuming and coupled with other PNA-specific issues (such as sequence constraints due to increased self-complementarity and decreased solubility in physiological buffers) this method might not be completely fit-for-purpose. As such, this section describes an

additional technique for the design of strand displacement reactions capable of interfacing D and L sequences.

In 2006, Hauser *et al* briefly described the $T_m$s of a handful of chimeric D/L-DNA duplexes (that is, half the duplex was D-DNA and half was L-DNA). These chimeric duplexes suffered a modest decrease in Tm but were otherwise stable[5]. On this basis, we devised an alternative to PNA-based heterochiral strand displacement systems using chimeric D and L-DNA strands, and two unique inversion gates were tested (Figure II-8 A and B). Both of these chimeric complexes each contain two strands composed of both D and L-DNA linked by a 3′-5′ phosphodiester linkage, while reaction B contains a third L-DNA strand making a 3-way junction (3WJ) motif. Every DNA strand in this scheme can be synthesized using standard phosphoramidite chemistry, rapidly speeding up the concept-to-lab adoption. Both reactions involve toehold-mediated nucleation of a D-nucleic acid input, followed by branch migration through the D-DNA region of the chimeric complex (domain 1 in figure II-8). The L-DNA hybridization of these complexes is relatively weak ($T_m$ < 30 ºC, domains 2 and 3 in Figure II-8) and they are predominately held together by their D-DNA stem, so after branch migration through domain 1 the L-DNA domains rapidly melt and a second toehold within domain 2 become activated to invade the reporter complex (Figure II-8 C). This reporter complex is nearly identical to the L-$R_{155}$ design described previously and will be denoted as L-$R_{155.2}$. It is important to note that in the case of these chimeric complexes, the L-DNA domains have no sequence dependence on the D-nucleic acid input, unlike the PNA/DNA inversion gate. This implies

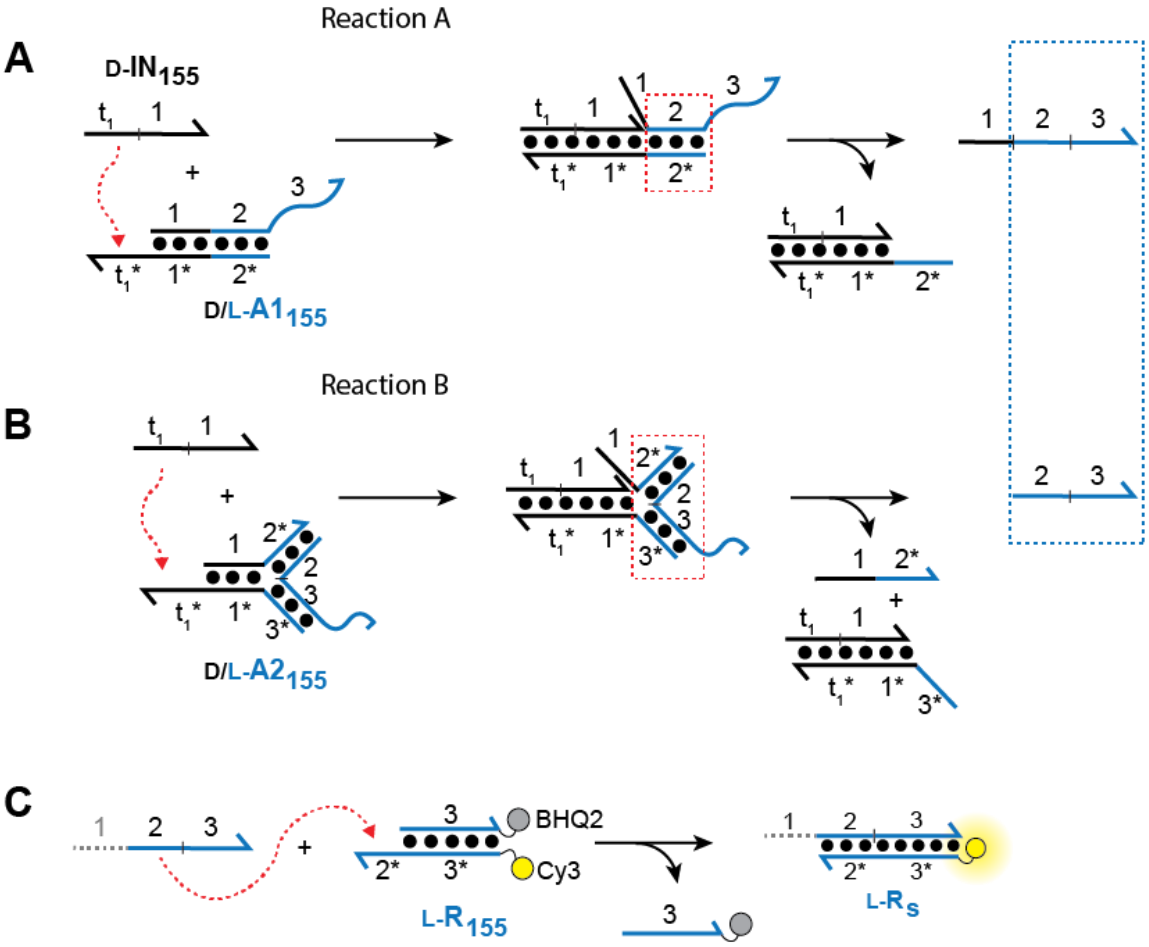



**Figure II-9. PNA-independent heterochiral strand displacement.** Schematic depiction of two PNA-independent heterochiral strand displacement designs based on chimeric D/L-DNA duplexes containing thermodynamically tuned L-DNA domains (red dashed boxes). (A) D-IN$_{155}$ binds to toehold $t_1$* and branch migrates through domain 1. L-DNA domain 2 is designed to have a $T_m$ below the reaction temperature, and spontaneously melts without the stability of domain 1. (B) A similar reaction containing a 3-way junction (3WJ) motif, strand displacement through domain 1 weakens domains 2 and 3 causing them to melt at the reaction temperature. (C) Reporter reaction with L-R$_{155.2}$ is initiated by either OUT strand. It is important to note that domains 1, 2 and 3 are the same in both reaction A and B.

that suitable L-DNA domains can be reused with an input of any sequence without constant re-validation.

Reactions were set up as described in section 2.3.5 and contained either 150 nM of the D-DNA input or 1 μM of the D-RNA input (Figure II-9). The rate constants of reactions initiated with the D-DNA input were calculated as described in section 2.1.1, yielding $1.5x10^4$ and $2.9x10^4$ $M^{-1}$ $s^{-1}$ for reactions A and B, respectively. Each of these heterochiral strand displacement reactions are approximately an order of magnitude faster than the previously described heterochiral strand displacement systems utilizing the PNA/DNA duplex L-$A_{155}$, but it is important to note that they each contain an 8-base toehold as opposed to the 6-base toeholds used previously. Interestingly, reactions with the D-RNA input were much slower than reactions with D-DNA input (apparent rate constants of $1.9x10^3$ and $4.2x10^3$ $M^{-1}$ $s^{-1}$ for reaction A and B, respectively), and required a large excess of the D-RNA input to push the displacement to completion. These results are apparently in contrast with literature data, which suggests that the hybridization kinetics of RNA onto DNA should be very similar to those of DNA onto DNA[23]. This is also the opposite trend of what we observed with the PNA/DNA system based on the L-$A_{155}$ duplex, where the heterochiral RNA input was faster than the heterochiral DNA input of the same sequence. Overall, the rates of strand displacement by a heterochiral RNA input from the PNA/DNA duplex L-$A_{155}$ and the chimeric D/L-DNA duplexes were very similar (rate constants on the order of ~$10^3$ $M^{-1}$ $s^{-1}$). The observed trends can likely be explained by the dramatically decreased $k_d$ of PNA/RNA hybridization, which can be ~200-fold lower than PNA/DNA interactions of the same sequence[16]. Additionally, the

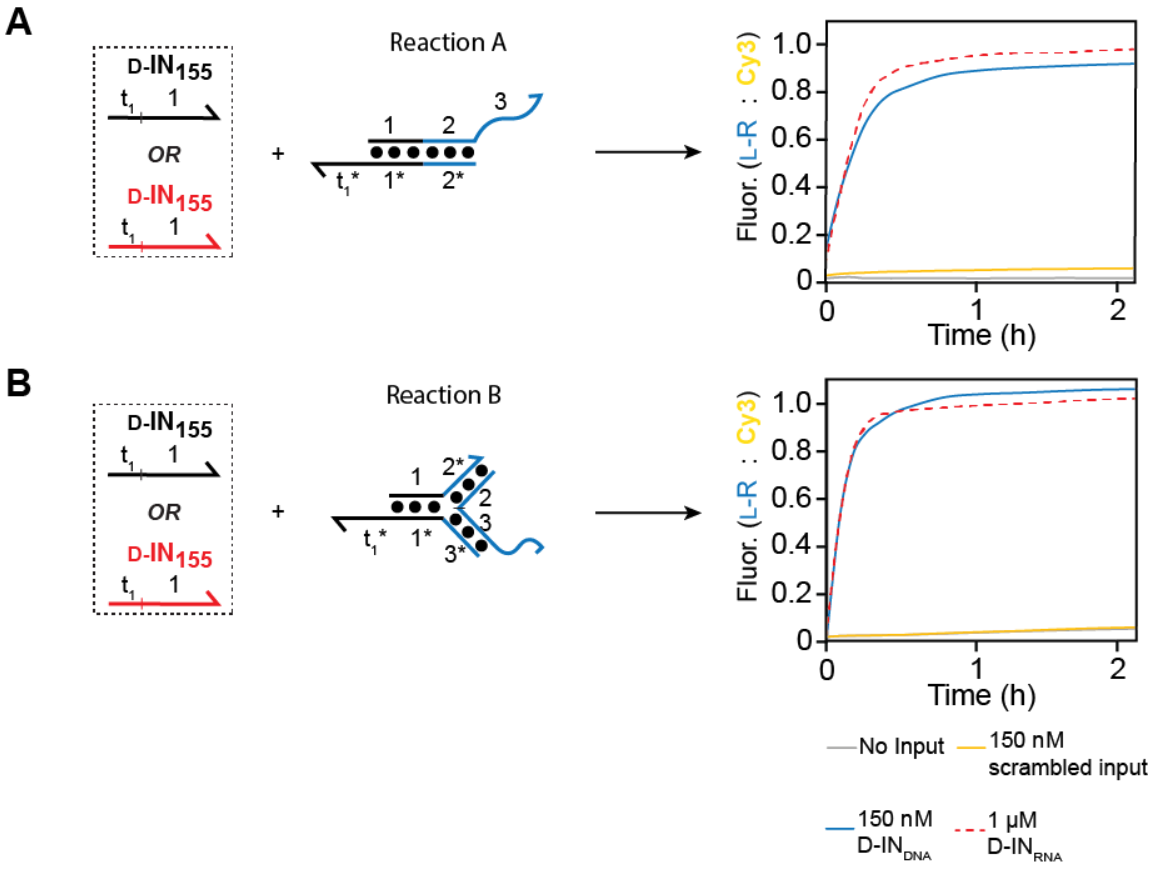



**Figure II-10. PNA-independent heterochiral strand displacement can detect either DNA or RNA inputs.** (A) Reaction A can be initiated by either DNA (black) or RNA (red) inputs. (B) Reaction B can be initiated by either DNA (black) or RNA (red) inputs. In either case, the RNA input needed to be higher concentration than the DNA input to react at a similar rate. This is likely due to secondary structure within the RNA input that is not present in the DNA input. For both reactions an input with a scrambled toehold (yellow trace), made by replacing $t_1$ with $t_1$* on the DNA input, did not generate signal.

online secondary structure prediction tool, NUPACK[10], predicts weak but not insignificant secondary structure within the miR-155 input strand. This may be harder to overcome when binding to a DNA toehold as opposed to a PNA toehold. Finally, since the goal of these designs is the eventual interaction with endogenous cellular nucleic acids, both D/L-complexes were tested with the D-RNA versions of the input strand (which is the same sequence as miR-155) with or without excess non-specific RNA from HeLa extract (Figure II-10). Excess non-specific RNA did not activate these circuits in the absence of the correct input.

## 2.2 Conclusion

In summary, a series of novel heterochiral strand displacement reactions were designed and tested for their ability to sequence-specifically convert a D or L-input strand to an output strand of the opposite chirality. Two different reaction designs were evaluated: One using a PNA/DNA heteroduplex as a chirality "inversion" gate, and the other using chimeric D/L-complexes designed to become destabilized and melt apart after successful invasion of a D-DNA branch migration domain. Both types of systems recognize a D-DNA or D-RNA input and autonomously convert it to an L-DNA output. Rate constants were determined for these reactions at 37 ºC, showing that they are slower than the reported rates for their all-DNA counterparts. This is especially true in the case of chiral inversion utilizing a PNA/DNA heteroduplex, as there seems to be a kinetic penalty when an input of one chirality invades a PNA heteroduplex of the opposite chirality. Future work will focus on identifying optimal reporter domains and carefully

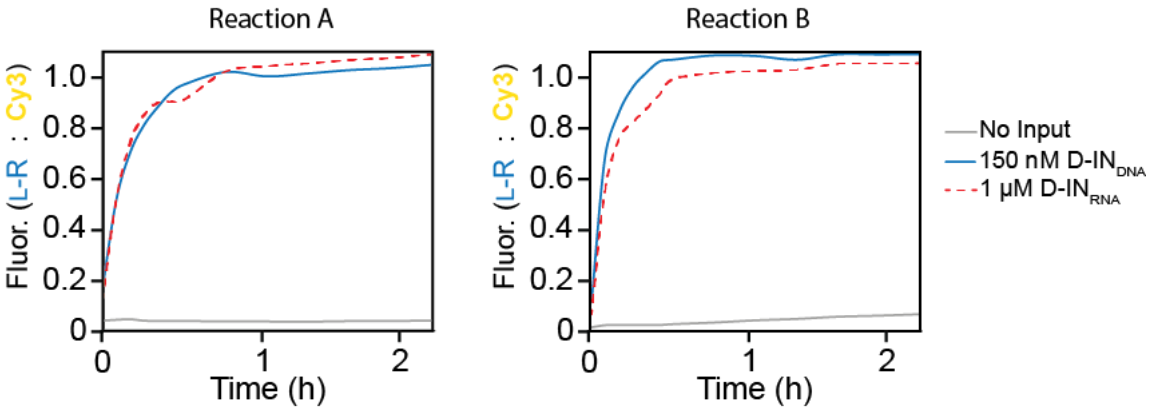



**Figure II-11. PNA-independent heterochiral strand displacement detects inputs in the presence of non-specific RNA.** Both Reaction A and B were inactive in the presence of 0.1 mg/mL total RNA from HeLa lysate and activated only in the presence of the correct inputs.

characterizing the effects different inputs have on strand displacement. I've described a first-of-its-kind thresholding gate, relying only on the chirality of its components rather that toehold modulation, that could be easily implemented to protect complex heterochiral strand displacement systems from unintended leak reactions *in vitro*. Further characterization of the kinetic properties of strand displacement from a PNA substrate are expected to yield valuable information informing the design of future generations of inversion gates.

## 2.3 Materials and Methods

### *2.3.1 DNA Design, synthesis and purification*

DNAs and RNAs were either purchased from Integrated DNA Technologies or prepared by solid-phase synthesis on an Expedite 8909 DNA/RNA synthesizer. DNA synthesis reagents, fluorophore phosphoramidites and D-nucleoside phosphoramidites were purchased from Glen Research, and L-nucleoside phosphoramidites were purchased from ChemGenes. Black Hole Quencher 2 (BHQ2) and Black Hole Quencher 3 (BHQ3) CPG resins were purchased from LGC Biosearch technologies. NHS ester version of the fluorescent dyes were purchased from Lumiprobe Life Science Solutions. PNAs were purchased from either PNA Bio Inc. or Panagene.

### *2.3.2 Sequence Design*

DNA sequences (Table A-2) for the strand displacement circuits depicted in this chapter were rationally designed and analyzed using NUPACK to ensure proper

hybridization and to decrease the risk of forming spurious secondary structures. PNA sequences were chosen based on complementarity to the desired RNA target, they were designed to minimize the number of purine residues to ensure adequate solubility, and regions >3 bases of self-complementarity were avoided. Since there is yet no tool for the prediction of PNA secondary structures, even though they are more prone to secondary structure formation than their natural DNA counterparts, all PNA sequences were input into NUPACK as RNA and analyzed for secondary structure. Since RNAs are prone to secondary structure formation this seemed like a good compromise, and all PNAs analyzed this way were well-behaved. PNA $T_m$s were approximated using the PNA Tool from PNA Bio Inc. All DNA melting temperatures were approximated using the IDT Oligo Analyzer tool which uses the nearest neighbor approximation.

### *2.3.3 Oligonucleotide purification*

After purchase from IDT, or after all modifications were finished in house, all oligonucleotides (D, L and chimeric D/L-strands) were purified by 20% denaturing polyacrylamide gel electrophoresis (PAGE; 19:1 acrylamide:bisacrylamide). Purified ONs were excised from the gel with a sterile razor blade and eluted overnight at room temperature (RT) in a buffer containing 200 mM NaCl, 10 mM EDTA, and 10 mM Tris pH 7.6. The solution was filtered to remove gel fragments and eluted oligonucleotides were precipitated with ethanol, resuspended in a small volume of ddH2O, and quantified by their absorbance at A260. All 3′ labeled oligonucleotides were synthesized using commercial CPG resins functionalized with the desired modification (e.g. BHQ2) and

purified as described above. Fluorescent dyes were either coupled directly to the 5′ end as their phosphoramidite using a 5-minute coupling protocol or attached post synthesis as an N-hydroxysuccinimide ester (NHS) via a 5′ amino modification installed at the time of synthesis. NHS ester conjugations reactions were performed by combining the unpurified amino modified oligonucleotide (~20 nmol per coupling reaction) with the desired dye NHS ester (5 mM final concentration) in 0.1 mL of 0.1 M sodium borate buffer (pH 8.5). Reaction mixtures were vortexed intermittently over 2 hours before being rocked gently overnight at RT. Excess fluorophore was removed by a NAP-5 Sephadex G-25 Column (GE Healthcare), and the labeled oligonucleotide was purified by 20% denaturing PAGE as described previously.

### *2.3.4 Preparation of duplex and 3WJ reaction components*

PNA/DNA duplexes were assembled via a hybridization titration approach in order to achieve an ideal 1:1 ratio of the corresponding strands. Here, one strand was held constant at 5 µM while the concentration of the second strand was varied across a narrow range around 5 µM. All hybridization mixtures contained the appropriate amount of each strand, 300 mM NaCl, 1 mM EDTA, 10 mM Tris pH 7.6 and were annealed by heating at 90 ºC for 3 minutes then cooled slowly to room temperature over 1 hour. The extent of hybridization was quantified by 20% native PAGE (19:1 acrylamide:bisacrylamide) after staining with SYBR Gold (Fisher Sci) and visualization on a Typhoon FLA 9500 (GE Healthcare; 488 laser line, 560 nm long pass filter). Only the PNA duplexes having a 1:1 ratio of strands were used further. Reporter sequences were designed such that the toehold

was on the fluorescently labeled strand, meaning the corresponding quencher strand could be used in slight excess without impacting circuit performance. As such, all reporters were annealed at the temperature and buffer conditions described previously with 10 µM of the fluorescent strand and 11 µM of the quencher strand.

The chimeric complexes were annealed at the temperature and buffer conditions as previously described, with 50 µM output strand (C1 and L-OUT) and 75 µM strands C2 (D/L-$A_1$) or C3+C4 (D/L-$A_2$), see Table A-2 for sequences. The complexes were further purified by 10% native PAGE to remove excess strands. Purified complexes were excised, crushed, and elution in a buffer containing 300 mM NaCl, 1 mM EDTA, 10 mM Tris pH 7.6 and eluted over 2 days at RT.

### *2.3.5 Monitoring of strand-displacement reactions by spectrofluorimetry*

Each Strand displacement reaction was monitored using a GloMax Discover multi-well plate reader (Promga Crop). All reactions using the PNA inversion gate contained 750 nM of the desired input (D-$IN_{DNA}$, L-$IN_{DNA}$, D-$IN_{RNA}$), 500 nM D or L-$A_X$, 1.5 µM L-$R_X$, 300 mM NaCl, 1 mM EDTA, and 10 mM Tris pH 7.6. Reactions were carried out in a 384-well microplate at a final volume of 30 µL and a temperature of 37 °C. Reactions were initiated by the addition of 10 µL of D or L-A duplex at the appropriate concentration. The fluorescence intensity of the D-reporter complex was monitored with excitation/emission wavelengths at 520 nm/580-640 nm (bandpass filter; Cy3), while the fluorescence intensity of the L-reporter complexes was monitored with

excitation/emission wavelengths at 627 nm/660-720 nm (bandpass filter; Cy5) or 520 nm/580-640 nm (bandpass filter; Cy3) depending on the system being studied.

All reactions using the chimeric (D/L) inversion gates contained 150 nM or 1 μM inputs (DNA or RNA, respectively), 100 nM either D/L-$A_1$ or D/L-$A_2$, and 300 nM L-$R_{155.2}$. These reactions were carried out in the same reaction buffer as described previously, initiated with 10 μL either D/L-$A_1$ or D/L-$A_2$, and carried out at 37 °C in a 384-well plate. Fluorescence was monitored with excitation/emission wavelengths at 520 nm/580-640 nm (bandpass filter; Cy3).

All strand displacement reactions were initiated to the signal from a pre-opened reporter representing the maximum achievable fluorescence using the following equation:

$$\boldsymbol{F_n = \frac{F - F_0}{F_c - F_0}}$$

Where $F_n$ is the normalized fluorescence intensity, F is the measured fluorescence at each time point, $F_0$ is the quenched fluorescence, and $F_c$ is the control fluorescence at each time a measurement was taken. This normalization equation allows for us to account for the loss in signal due to photobleaching and enables the direct comparison of the different fluorophores used in this study.

### *2.3.6 Monitoring of strand-displacement kinetics by spectrofluorimetry*

For reactions involving PNA/DNA heteroduplex L-$A_X$ (Figures II-3, II-4, II-5 and II-7), we observed a substantial difference in the rate of strand-displacement based on the chirality of the input strand. We determined the rate constant for each of these strand displacement reactions as previously reported[17]. Briefly, the strand displacement reactions

were prepared and initiated as described in the previous section and monitored by spectrofluorimetry. Under these conditions, kinetics of reporter opening does not limit the overall reaction rate. All kinetics reactions were performed with 1.5 equivalents of the desired input strand relative to the inversion gate in order to ensure complete displacement of the output strand. The fluorescence curves obtained were fit using an equation derived from the second order rate law with respect to the input strand and the inversion gate. Due the stability of the waste duplexes, all reverse reactions were considered negligible. Strand displacement kinetics for reactions involving the chimeric complexes D/L-$A_1$ and D/L-$A_2$ were calculated in the same fashion.

## 2.4 References


1. Garbesi, A., et al., L-DNA as potential antimessenger oligonucleotides: a reassessment. *Nucleic Acids Res.* **1993,** *21* (18), 4159-4165.

2. Egholm, M., et al., Peptide nucleic acids (PNA). Oligonucleotide analogs with an achiral peptide backbone. *J. Am. Chem. Soc.* **1992,** *114* (5), 1895-1897.

3. Egholm, M., et al., Recognition of guanine and adenine in DNA by cytosine and thymine containing peptide nucleic acids (PNA). *J. Am. Chem. Soc.* **1992,** *114* (24), 9677-9678.

4. Egholm, M., et al., PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules. *Nature* **1993,** *365* (6446), 566-568.

5. Hauser, N. C., et al., Utilising the left-helical conformation of L-DNA for analysing different marker types on a single universal microarray platform. *Nucleic acids Res.* **2006,** *34* (18), 5101-5111.

6. Hoehlig, K., et al., Stereospecificity of oligonucleotide interactions revisited: no evidence for heterochiral hybridization and ribozyme/DNAzyme activity. *PLoS One* **2015,** *10* (2), e0115328.

7. Kabza, A. M., et al., Heterochiral DNA Strand-Displacement Circuits. *J. Am. Chem. Soc.* **2017,** *139* (49), 17715-17718.

8. Mattiske, S., et al., The Oncogenic Role of miR-155 in Breast Cancer. *Cancer Epidemiol., Biomarkers Prev.* **2012,** *21* (8), 1236.

9. Kim, S., et al., microRNA-155 positively regulates glucose metabolism via PIK3R1-FOXO3a-cMYC axis in breast cancer. *Oncogene* **2018,** *37* (22), 2982-2991.

10. Zadeh, J. N., et al., NUPACK: Analysis and design of nucleic acid systems. *J. Comput. Chem.* **2011,** *32* (1), 170-173.

11. Srinivas, N., et al., On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.* **2013,** *41* (22), 10641-10658.

12. Isaksson, J.; Chattopadhyaya, J., A Uniform Mechanism Correlating Dangling-End Stabilization and Stacking Geometry. *Biochemistry* **2005,** *44* (14), 5390-5401.

13. Morrow, S. M., et al., Transmission of chirality through space and across length scales. *Nat. Nanotechnol.* **2017,** *12*, 410.

14. Wittung, P., et al., DNA-like double helix formed by peptide nucleic acid. *Nature* **1994,** *368* (6471), 561-563.

15. Sforza, S., et al., Chiral Peptide Nucleic Acids (PNAs): Helix Handedness and DNA Recognition. *Eur. J. Org. Chem.* **1999,** *1999* (1), 197-204.

16. Jensen, K. K., et al., Kinetics for Hybridization of Peptide Nucleic Acids (PNA) with DNA and RNA Studied with the BIAcore Technique. *Biochemistry* **1997,** *36* (16), 5072-5077.

17. Olson, X., et al., Kinetics of DNA Strand Displacement Systems with Locked Nucleic Acids. *J. Phys. Chem. B* **2017,** *121* (12), 2594-2602.

18. Zhang, D. Y.; Winfree, E., Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.* **2009,** *131* (47), 17303-17314.

19. Yurke, B.; Mills, A. P., Using DNA to Power Nanostructures. *Genetic Programming and Evolvable Machines* **2003,** *4* (2), 111-122.

20. Sheedy, P.; Medarova, Z. The fundamental role of miR-10b in metastatic cancer *Am J Cancer Res* [Online], 2018, p. 1674-1688. PubMed. http://europepmc.org/abstract/MED/30323962 (accessed 2018).

21. Hart, P. C., et al., MnSOD upregulation sustains the Warburg effect via mitochondrial ROS and AMPK-dependent signalling in cancer. *Nat. Commun.* **2015,** *6* (1), 6053.

22. Seelig, G., et al., Enzyme-Free Nucleic Acid Logic Circuits. *Science* **2006,** *314* (5805), 1585.

23. Rauzan, B., et al., Kinetics and Thermodynamics of DNA, RNA, and Hybrid Duplex Formation. *Biochemistry* **2013,** *52* (5), 765-772.

# CHAPTER III

# STUDYING THE PERFORMANCE OF HETEROCHIRAL STRAND DISPLACEMENT CIRCUITS IN LIVING CELLS

Molecular computation has advanced considerably since Adleman first encoded a solution to the Hamiltonian path problem in DNA[1]. The ability to routinely label synthetic DNA strands with fluorescent dyes and quencher molecules enabled the synthesis of "digital" DNA nanodevices that could be turned ON or OFF with high specificity thanks to the programmable nature of Watson – Crick (WC) base pairing. More recently, the development of toehold-mediated strand displacement reactions has dramatically increased the scope of DNA computation and molecular recognition. These developments lead to a surge in complexity for DNA computation in vitro[2-4], and spurred efforts to harness the power of DNA computation in live cells[5-6]. However, the biological environment inside a cell is vastly different than the carefully controlled *in vitro* environments where these systems usually perform. Generally, there are 4 primary considerations when designing a DNA device to perform in the cellular environment: 1) the DNA device needs to be delivered across the cell membrane, 2) it needs to be stable in the cellular milieu (resistant to nuclease degradation and off-target interactions with other biomacromolecules), 3) it needs to be able to perform in an environment with a high degree of molecular crowding, and 4) it needs to be non-toxic/non-immunogenic.

Unsurprisingly, L-DNA seems to be an ideal modification for this application. L-DNA is known to be stable against nucleolytic degradation with little observed toxicity or immunogenicity[7-8]. However, the use of L-DNA does not circumvent the primary concern

facing nucleic acid – based probes and therapeutics: the cell membrane (see sections 1.4.2 and 1.5.2). This chapter focuses on the screening of transfection reagents and sample conditions for cellular delivery of heterochiral strand displacement circuits based on the PNA/DNA design described in chapter 2. Importantly, these results highlight some of the limitations of liposome and polymer – based transfection reagents for the delivery of small DNA components. The activation of single input and multi input heterochiral strand displacement systems is described in live cells and compared to the activation of systems with the same sequence containing D-DNA components. Additionally, a version of the miR-155 system in which all DNA and PNA components were replaced with 2′-O-methyl-RNA (2′-OMe-RNA) bases was tested as a comparison. This is one of the most common modifications currently used to increase the stability of DNA components inside live cells, and we expected it to behave similarly to the L-DNA system. The results in this chapter establish the importance of reporter stability and the simplicity of a "plug-and-play" L-DNA approach as opposed to other traditional DNA modifications.

## 3.1 Results

### *3.1.1 Screening transfection reagents and the in vitro optimization of transfection conditions*

Cationic transfection reagents are some of the most common tools for the delivery of nucleic acids into living cells[9-10]. These reagents represent effective delivery vehicles for many cell types, and their efficiency represents the gold standard for non-viral nucleic acid delivery. However, there are dozens of cationic transfection agents available

commercially, with multiple formulations that are optimized for a variety of nucleic acid cargo (i.e. siRNA, mRNA, plasmids, etc.). This is further complicated by the relatively young field of biocompatible DNA nanodevices. The components used in these devices tend to be relatively short, similar in length to siRNA, which may complicate the complexation of negatively charged DNA with the positively charged transfection reagent. To that end, a panel of common transfection reagents were tested to evaluate their ability to deliver these circuit components. In this assay, a short single stranded L-DNA labeled with the Cy3 dye was incubated with 3 different amounts of the indicated transfection reagent and allowed to form complexes according to the manufacturer's protocol. Many of these transfection reagents may have had a higher transfection efficiency with a double stranded DNA probe, as it would represent a more canonical nucleic acid cargo (e.g. siRNA or double stranded DNA). However, in many of the following experiments cells are transfected with single-stranded D or L-DNA inputs, and the PNA/DNA inversion gates (D or L-$A_X$, where X denotes which system is being used) only have the charge character of single stranded DNA due to the uncharged nature of the PNA backbone. On this basis, choosing a transfection reagent capable of delivering single-stranded DNA was determined to be more useful in the context of this experimental design. Transfection reagents were screened with an L-DNA cargo due to its increased stability, as degradation of a single-stranded DNA might lead to non-specific dye uptake within cells[11]. After complexation, DNAs were added to HeLa cells and incubated for two hours before being imaged by fluorescence microscopy (Figure III-1). While 4 of the tested transfection reagents show moderate delivery of the short single-stranded DNA, Lipofectamine

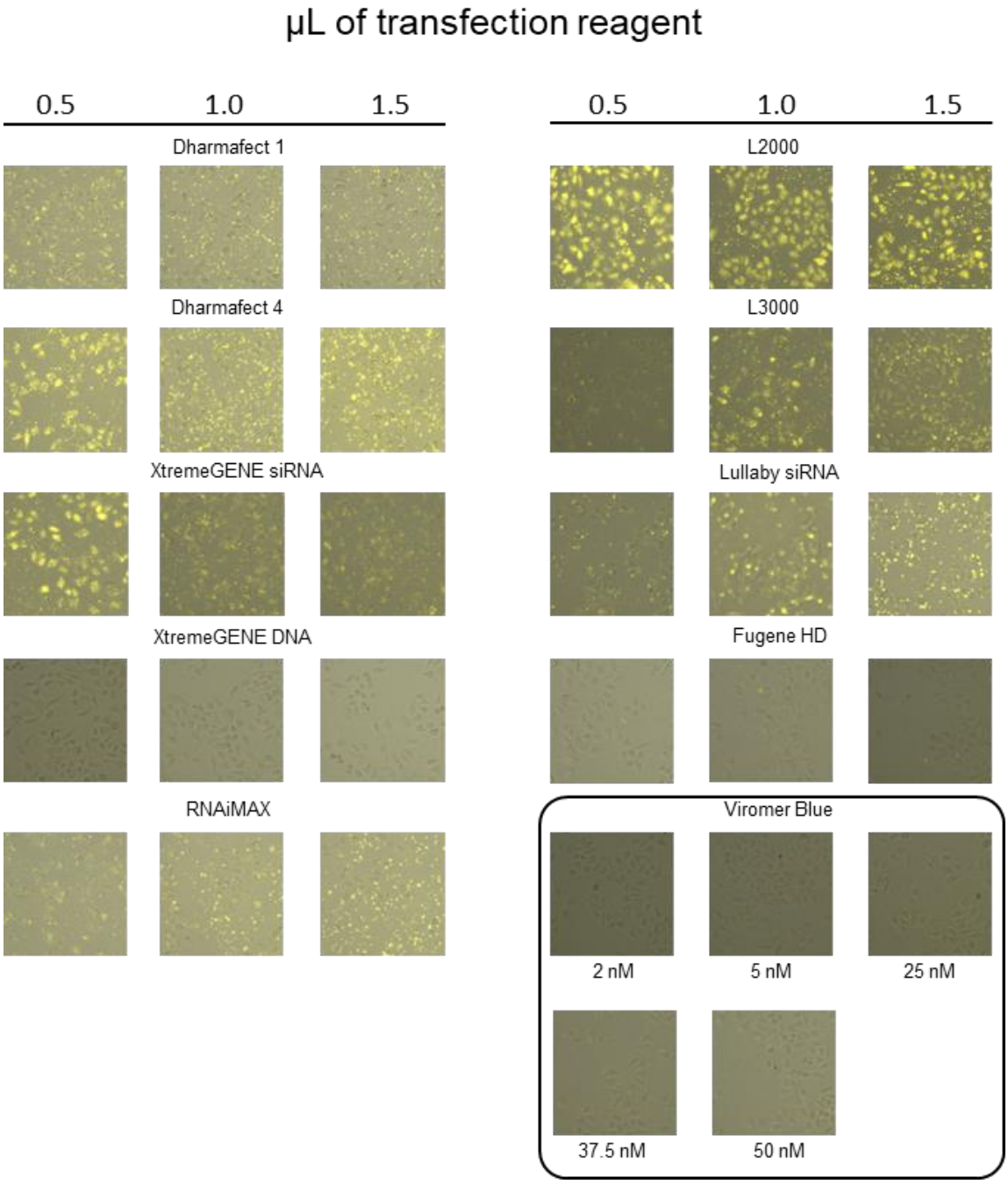



**Figure III-1. Transfection reagent screen.** A panel of transfection reagents were tested for their ability to deliver a short, single-stranded L-DNA into cells. Transfection complexes were prepared following the manufacturer's protocol using either 0.5, 1.0, or 1.5 µL per 1 pmol of DNA. An expanded screen of the Viromer Blue transfection reagent was performed as recommended by the manufacturer.

2000™ (L2000) consistently resulted in the highest delivery. With this in mind, L2000 was used for the remainder of the experiments described in this chapter.

To minimize the complexity of these experiments, it would be optimal to deliver all of the desired circuit components at the same time. However, since we wanted to characterize these reactions *inside* the cell, it was important to confirm that these strand displacements wouldn't occur when the components were mixed after complexation but before delivery. This was tested with the modified version of the miR-155 (D or L-$R_{155.2}$) system, the miR-10b system, and the L-AND gate described in the previous chapter (Figure III-2 A). Literature suggested that L2000 would inhibit the interaction of individual components *in vitro*[12-13] (Figure III-2 B), and this was verified by spectrofluorimetry. Components were complexed separately in L2000, incubated for 20 minutes, then mixed together immediately before analysis (Figure III-3 A). Briefly, reactions were prepared in OptiMEM at 2X the final concentration, split in half, and diluted in OptiMEM with or without the indicated amount of L2000 to 1X. Multiple concentrations of L2000 were tested based on the initial transfection reagent screen (data not shown). It was determined that for the single input heterochiral strand displacement systems (miR-155 and miR-10b cascades) containing 25 nM L-$R_X$, 50 nM of L-$A_X$ and 50 nM of the indicated input (D or L-$IN_X$), 0.4 µL L2000 per component was sufficient to inhibit strand displacement for up to 1.5 hr (Figure III-3 C, 155 and 10b) when reaction components were mixed in a test tube. The multicomponent AND gate was also prepared in 3 DNA – L2000 mixtures: the inputs (2 single stranded DNAs), the inversion gates (2 PNA/DNA heteroduplexes) and the AND reporter (twice as long as the single input

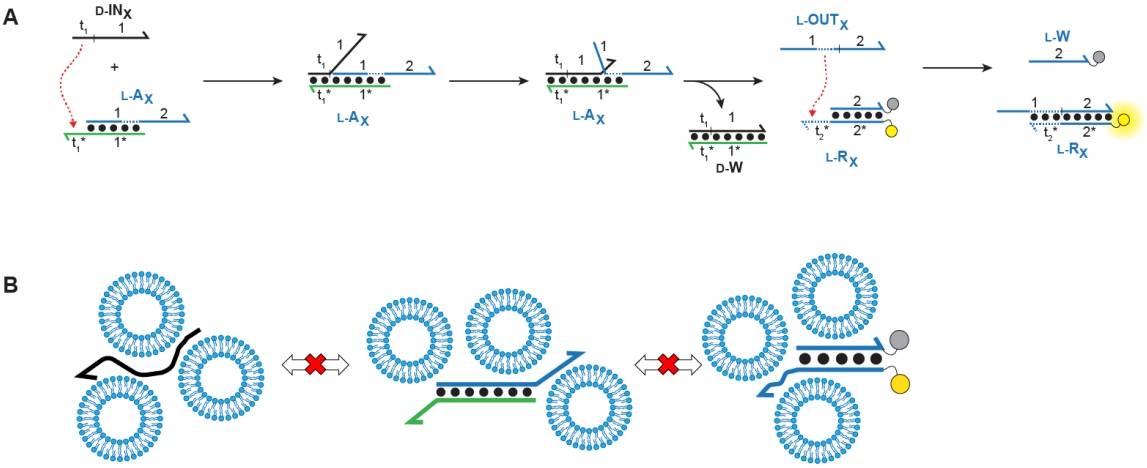



**Figure III-2. Schematic of strand displacement.** (A) Heterochiral strand displacement scheme. (B) L2000 is expected to package individual components into lipoplexes, and these lipoplexes should prevent circuit activation until the strands are released into cells.

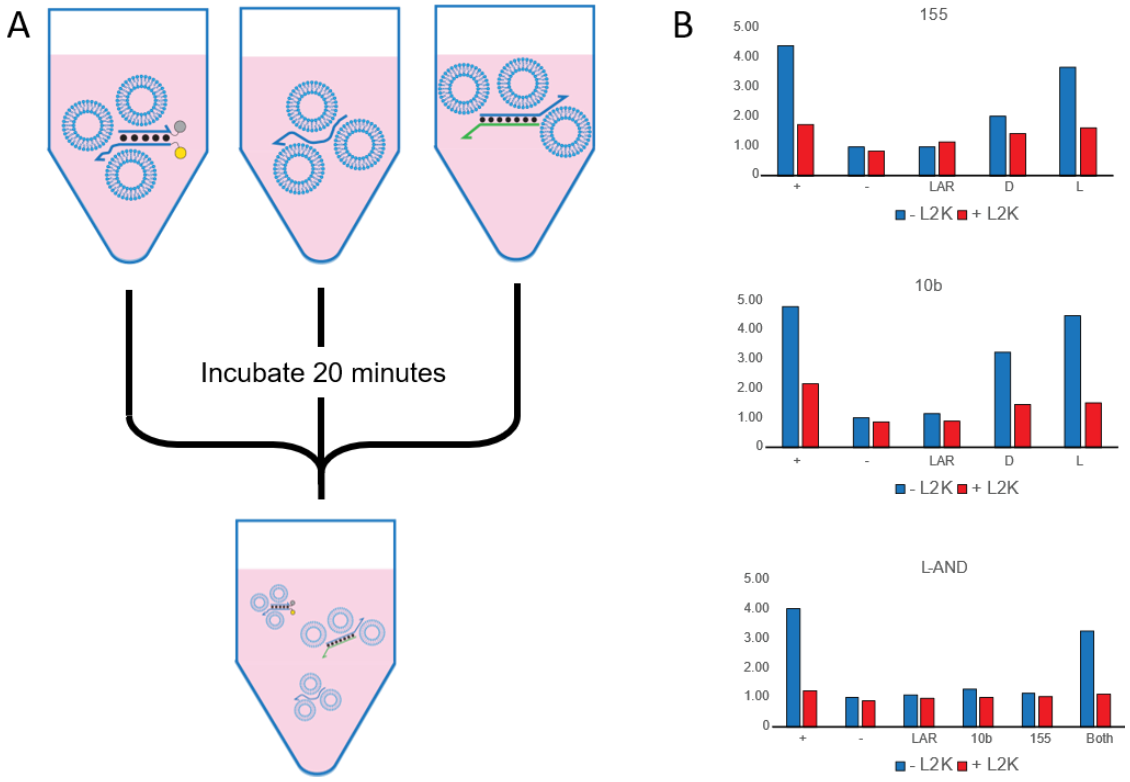



**Figure III-3. Circuit suppression by L2000.** (A) Components were diluted individually into optiMEM with or without L2000 to a final concentration of 75 nM L-$R_X$, 150 nM L-$A_X$, and 150 nM D or L-$IN_X$. Solutions were incubated for 20 minutes to form lipoplexes, then mixed, added to a 384 well plate, and the following displacements were monitored by spectrofluoimetry: Positive control (+, L-$OUT_X$ and L-$R_X$), negative control (L-$R_X$ alone), AR (L-$A_X$ and L-$R_X$ mixed), D (D-$IN_X$, L-$A_X$ and L-$R_X$) and L (L-$IN_X$, L-$A_X$ and L-$R_X$). In these displacement X refers to the system being tested (miR-155, miR-10b, or L-AND). (B) After 1.5 hours the activation of each mixture with or without L2000 was compared. L2000 generally suppressed activation to < 2-fold the negative control.

reporters). For each of these mixtures, 0.8 μL of L2000 was required to prevent component interaction over the same time frame (Figure III-3 B, L-AND). This is consistent with the overall charge of each component mixture increasing approximately 2-fold.

### *3.1.2 Determining the stability of L, D and 2′-O-methyl reporters in live cells*

The reporter module is one of the most important components in the canonical DNA strand displacement system, and this is even more true in the complex biological environment of a cell. Since the reporter is responsible for generating the output signal after the detection of a strand of interest, reporter activation (ON) and stability (OFF) directly affect the signal-to-background ratio of these sensors. As such, it was important to test the activation of each of the 3 different reporter constructs, based on the $R_{155.2}$ sequences, both *in vitro* and in live cells. The three reporters used in these experiments: an all L-DNA reporter, an all D-DNA reporter, and a 2′-OMe RNA reporter (Figure III-4 A) were the same sequence, differing only in their backbone modifications. Each reporter was tested *in vitro* by activation with their respective OUT strand to ensure that they reached similar maximum signals. As expected, the D-DNA and L-DNA reporters behaved similarly, rapidly reaching their maximum signals after initiation with a 10-fold excess of their respective OUT strand. The 2′-OMe-RNA reporter, however, was not fully activated under these conditions, although a larger excess of input (~100-fold) allowed this reporter to reach maximum signal in the same time frame (Figure III-4 B). While the DNA $OUT_{155}$ strands were predicted to have very little secondary structure in the online secondary structure tool NUPACK[14], if the same sequence is treated as RNA (which will more

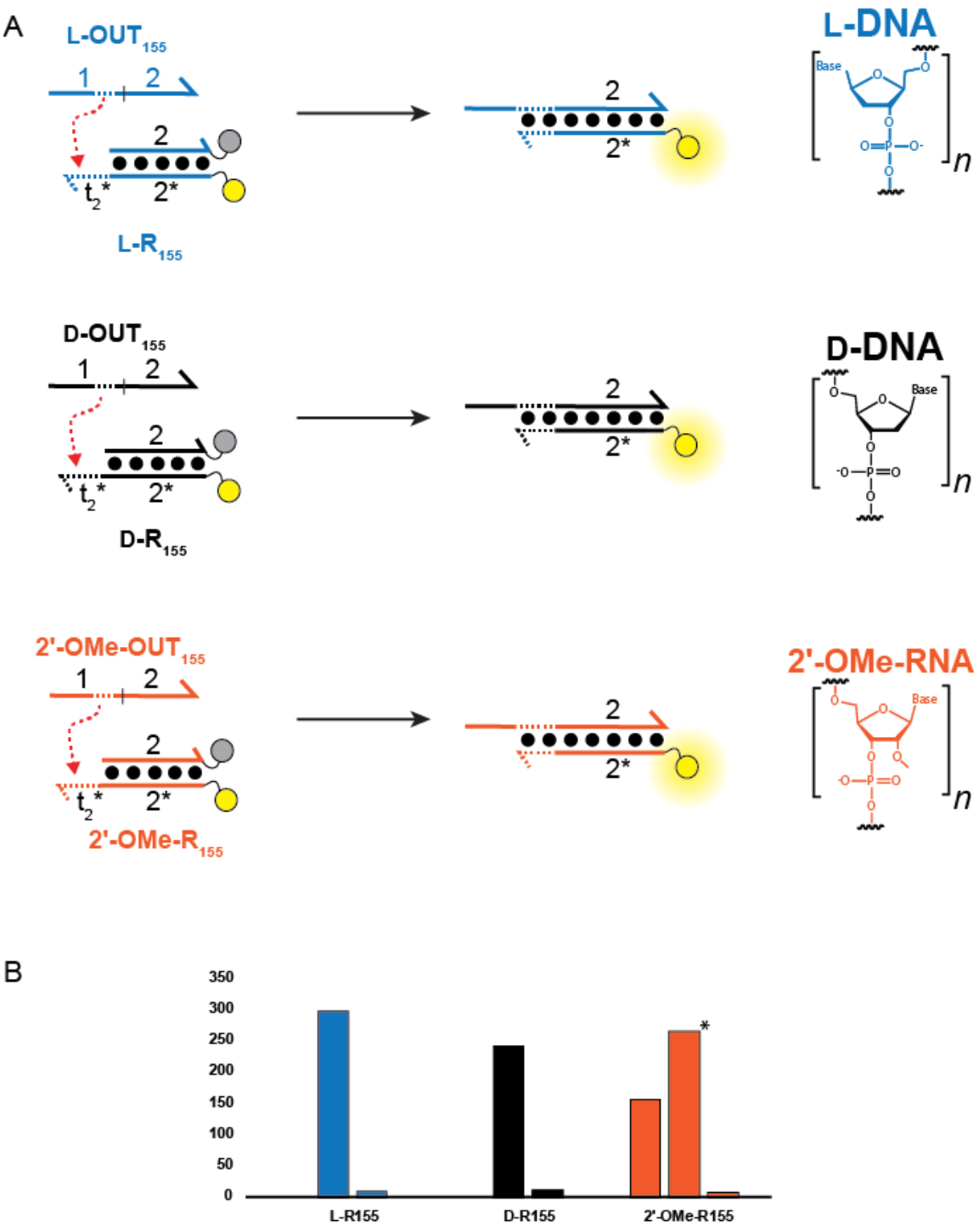



**Figure III-4. Reporters tested in this section.** (A) Schematic of the activation of L-DNA (blue), D-DNA (black) and 2′-OMe (orange) reporters described in this section. (B) Activation of each reporter by 10-fold excess $OUT_{155}$ strand of the indicated chirality, and the background of quenched reporters. The asterisk (*) indicates 100-fold excess $OUT_{155}$.

accurately represent strands containing the 2′-OMe modification) it is predicted to form a stable hairpin. This type of secondary structure is known to impact strand displacement and highlights the utility of L-DNA since its thermodynamic parameters should exactly mimic those of D-DNA.

After analysis *in vitro,* these reporters were transfected into HeLa cells to see if they would behave similarly in the complex cellular environment. Each reporter and its OUT strand (D-$R_{155.2}$ + D-$OUT_{155}$, L-$R_{155.2}$ + L-$OUT_{155}$ and 2′-OMe-$R_{155.2}$ + 2′OMe-$OUT_{155}$) were incubated separately in L2000 for 20 minutes to allow the formation of lipoplexes. After incubation, each reporter was mixed with its $OUT_{155}$ strand or OptiMEM alone ('+' or '-', Figure III-5) before being added to HeLa cells and incubated for 1.5 hours. Lipoplexes were then removed from the wells by aspiration and replaced with fresh DMEM, and the cells were imaged at 2, 4, 6 and 24 hr after transfection (Figure III-5 A, B, C and D, respectively). After imaging, cells were trypsinized and analyzed by flow cytometry to quantify reporter activation and background. The bar graphs shown in Figure III-5 represent the geometric means of the Cy3 signal after subtraction of the background fluorescence from a non-transfected population of cells (see section3.3.5 for more details). Each bar represents the fold enhancement of the indicated cell population after normalization to the mean intensity of L-$R_{155.2}$ at the two-hour timepoint. These results demonstrate the sustained activation of L-$R_{155.2}$ after cotransfection with L-$OUT_{155}$, its relative stability in the absence of the L-$OUT_{155}$ strand, the rapid degradation of D-$R_{155.2}$, and the poor activation of 2′-OMe-$R_{155.2}$ (mimicking what was demonstrated in Figure

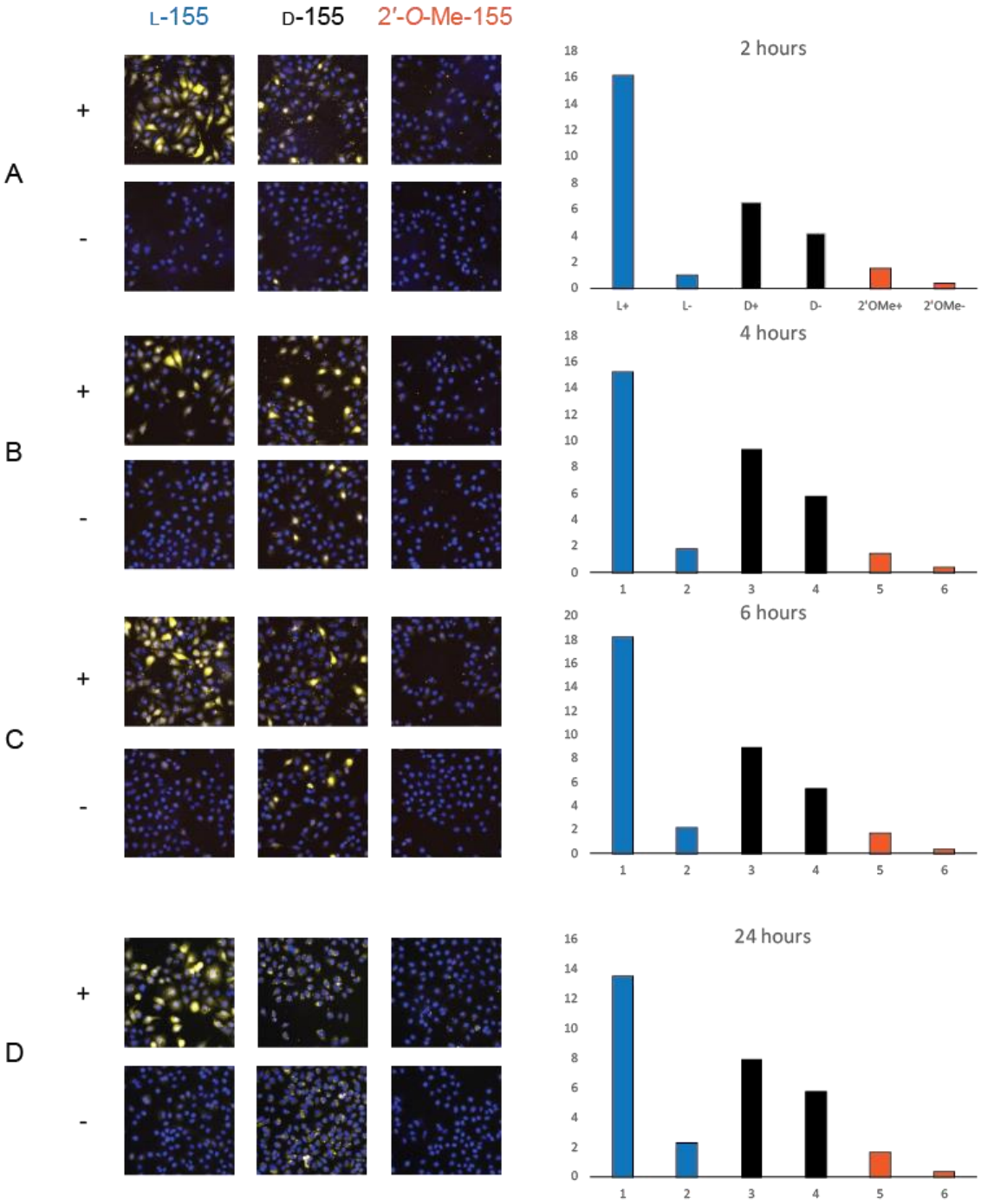



**Figure III-5. Time course of reporter activation in live cells.** Representative images showing the activation of the 3 different reporters at 2 hours (A), 4 hours (B), 6 hours (C) and 24 hours (D). Bar graphs represent flow cytometry data from each experimental well normalized to the quenched L-$R_{155}$ at the 2-hour time point.

III-4 B). The results shown in Figure III-5 clearly demonstrate the utility of L-DNA as a modification in biostable DNA devices.

*3.1.3 Single component heterochiral strand displacement systems in live cells*

After the three reporters were characterized individually, we next wanted to test if the full heterochiral strand displacement cascade (see Figure III-2 A for the general strand displacement scheme) could be activated inside live HeLa cells. Having previously shown that L2000 prevents the interaction of all 3 gate components *in vitro* for up to 1.5 hr, all components for the indicated cascade were co-transfected in this manner. Based on the data indicating the L-reporter has increased stability and activation compared to the D-DNA and 2′-OMe-RNA reporters, we hypothesized that the heterochiral strand displacement systems would have better intracellular performance than either of the D-nucleic acid strand displacement systems. This hypothesis was tested using variants of the miR-155 based strand displacement system (Figure III-6). Similar results were obtained when transfecting the L vs. D-DNA variants of the miR-10b system (Figure III-7). Briefly, each component was mixed with 0.4 µL L2000 in optiMEM at a final volume of 40 µL and incubated for 20 minutes to form lipoplexes. The components were then mixed as indicated in table III-1 (D or L-miR-155 and D or L-miR-10b systems) or table III-2 (2′-O-Me-155 system), then added to HeLa cells for 1.5 hr before the transfection mixture was removed by aspiration and replaced with fresh media. After an additional 4.5 hours (for a total of 6 hours) cells were imaged by fluorescence microscopy, trypsinized and fluorescence was quantified by flow cytometry (Figures III-6 B and III-7 B). In these

| | Experimental Wells | | | | | | |
|---|---|---|---|---|---|---|---|
| **DNA strand** | **+** | **-** | **No input (1)** | **D-scramble (2)** | **D-input (3)** | **L-scramble (4)** | **L-input (5)** |
| D or L-$R_X$ | + | + | + | + | + | + | + |
| D or L-$OUT_X$ | + | - | - | - | - | - | - |
| D or L-$A_X$ | - | - | + | + | + | + | + |
| D-scramble | - | - | - | + | - | - | - |
| D-input | - | - | - | - | + | - | - |
| L-scramble | - | - | - | - | - | + | - |
| L-input | - | - | | - | - | - | + |

**Table III-1.** Combinations of components mixed for experiments with the miR-155 and miR-10b systems.

| | Experimental Wells | | | | |
|---|---|---|---|---|---|
| **DNA strand** | **+** | **-** | **No input (1)** | **DNA input (2)** | **2′-OMe input (3)** |
| 2′-OMe-$R_{155}$ | + | + | + | + | + |
| 2′-OMe-$OUT_{155}$ | + | - | - | - | - |
| 2′-OMe-$A_{155}$ | - | - | + | + | + |
| D-DNA input | - | - | - | + | - |
| 2′-OMe input | - | - | - | - | + |

**Table III-2.** Components mixed for experiments with the 2′-OMe miR-155 system.

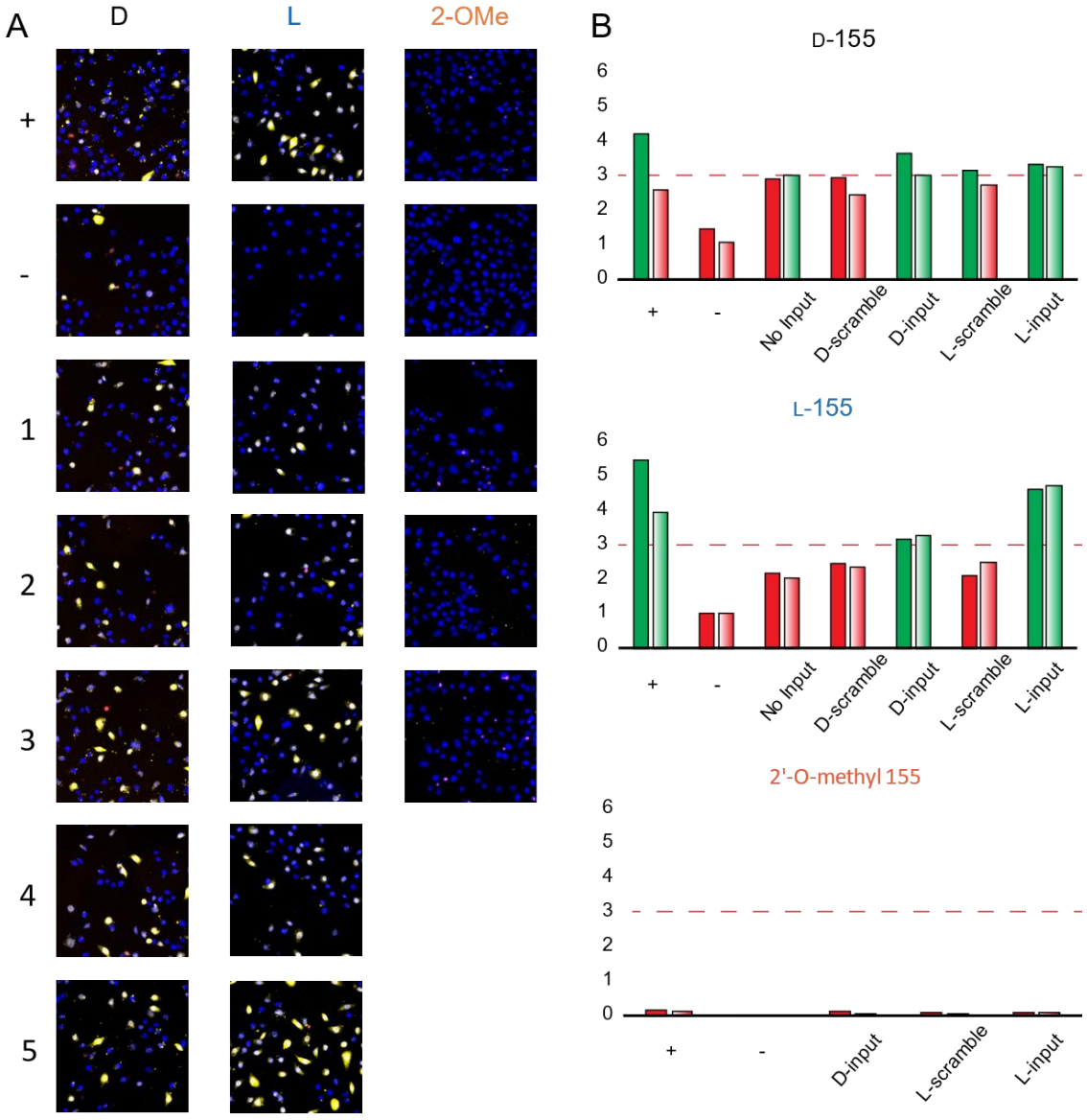



**Figure III-6. Activation of full miR-155 systems in live cells.** (A) Representative images of cells transfected with the D-155 system components, L-155 system components or the 2′-OMe-155 system components. Labels correspond to mixtures indicated in Tables III-1 and III-2. (B) Quantification of the activation of each circuit in live cells. In each case, solid bars and gradient bars indicate biological duplicates. Bars were generated by normalizing the mean fluorescence of each population to a quenched L-$R_{155}$ sample run on the same day.

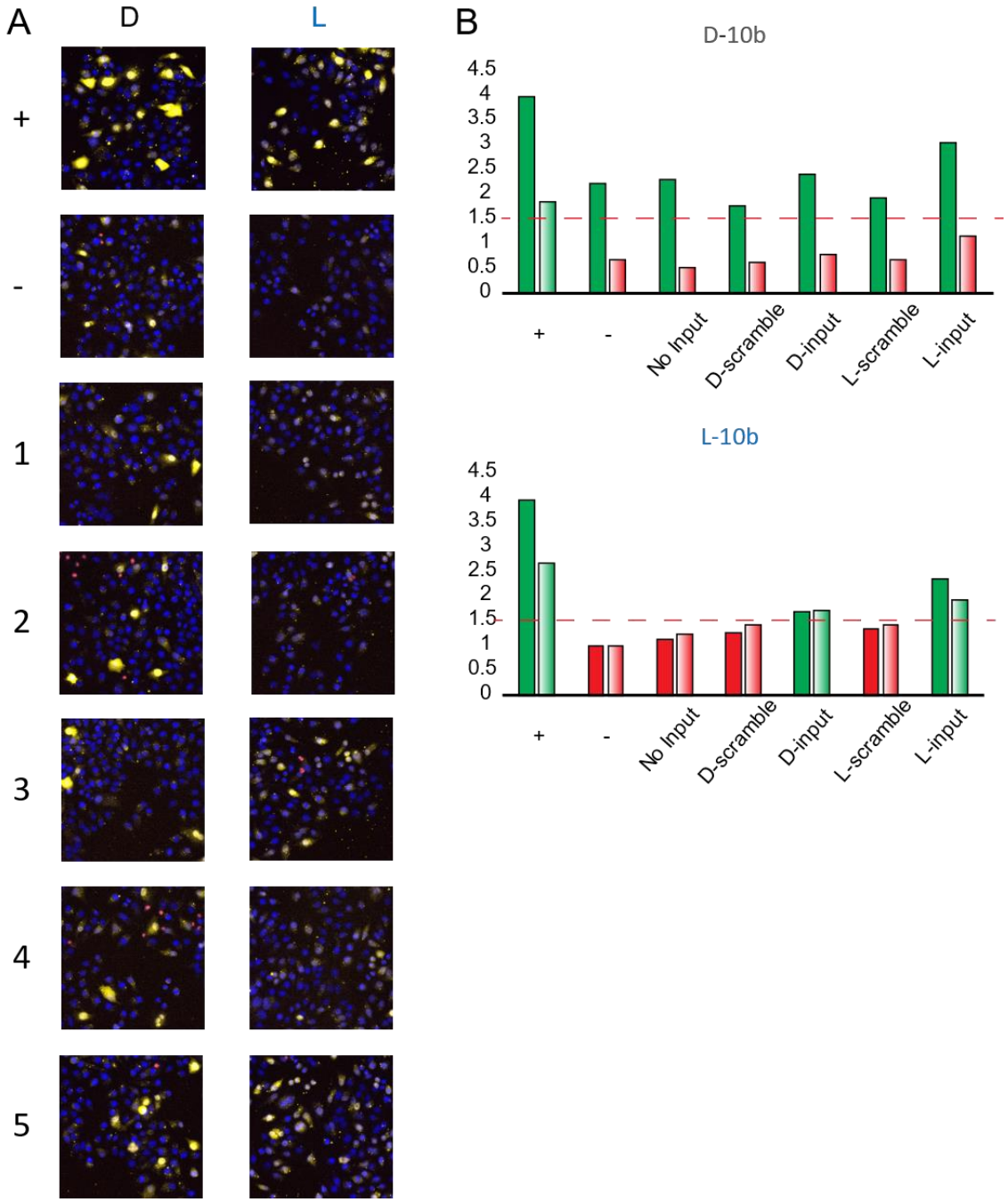



**Figure III-7. Activation of full miR-10b systems in live cells.** (A) Representative images of cells transfected with the D-10b system components or L-10b components. Labels correspond to mixtures indicated in Table III-1. (B) Quantification of the activation of each circuit in live cells. In each case, solid bars and gradient bars indicate biological duplicates. Bars were generated by normalizing the mean fluorescence of each population to a quenched L-$R_{10b}$ sample run on the same day

experiments it was important to verify that fluorescent signal is generated specifically due to the proposed toehold mediated strand displacement, so D and L-inputs containing a scrambled toehold were designed and used as additional controls (see table A-3 in appendix A for sequences). In this way, even though the branch migration domain is the same sequence the scrambled input should not displace the output strand from the PNA. Interestingly, in the miR-155 heterochiral strand displacement system these scrambled inputs activated the cascade to nearly identical levels as the inputs with the correct toehold in live cells (data not shown). Based on this observation, the inputs were truncated by two nucleotides at the 3′ end, introducing a "clamp" at the end of the L-$A_{155}$ branch migration domain. While this change impacted the maximum activation of heterochiral strand displacement in cells, it reduced toehold-independent activation to near background levels with respect to group 1 in Figures III-6. The full-length scrambled inputs were nearly inactive *in vitro*, highlighting the unique effects imposed by the cellular environment on DNA strand displacement systems.

While the results discussed in this section demonstrate the same trends as their *in vitro* counterparts (see section 2.1.1), there were a number of issues encountered in live cells. While both D-DNA systems (miR-155 and miR-10b) were consistently poor, as expected, the extent of degradation seemed surprisingly variable even between experiments performed on the same day (see D-10b, Figure III-7 B). The L-DNA systems, in contrast, were much more consistent in general. However, it is concerning that the leak of the miR-155 system is much higher in cells than *in vitro* experiments would predict (Figure III-6 B, L-155 no input experiment). While the scrambled controls used in these

experiments effectively suppressed strand displacement, they cannot reduce signal further than the background leak reaction. In contrast, the miR-10b system (L-10b, Figure III-7 B) demonstrated consistently lower activation than the miR-155 system even though the miR-10b system is much faster *in vitro*. The miR-10b and miR-155 experiments were performed at substantially different passage numbers, so it's certainly possible that this may manifest as a difference in transfection efficiency between these experiments. Under these experimental conditions, the transfection efficiency of a short single-stranded L-DNA is ~50% in my hands (data not shown), so anything that might adversely affect this (i.e. passage number or cell confluency) will likely have a large impact on the population of cells that receives all three of the individually packaged components. It is apparent from the data presented that future designs should take steps to mitigate leak between duplex A and the reporter, but the most beneficial change would likely come from the adoption or identification of a more robust delivery vehicle.

### *3.1.4 Multi-component heterochiral strand displacement systems in live cells*

The potential utility of biostable DNA computation is greatly increased by the relatively straight-forward design of higher order, multi-component systems with L-DNA. *In vitro*, such systems are able to respond to multiple single stranded DNA or RNA inputs and, using defined Boolean logic operations, generate a single DNA output based on a pre-determined computational path[2]. In a cellular setting the release of a DNA output strand can act as a sensor for specific combinations of RNA inputs that may help define a particular cell type or disease state. This represents a potentially attractive alternative to

traditional multiplexing techniques, which rely on the use of multiple spectrally resolved fluorophores to identify each additional target. To this end, the 2 – input heterochiral strand displacement system described in section 2.1.1 was transfected into cells to monitor reporter stability and circuit activation of a multi-input L-DNA sensor (Figure III-8).

Cells were transfected with each group of components as described previously (Table III-2), imaged by fluorescence microscopy and quantified by flow cytometry (Figure III-9). L-AND gate leak was assessed in the presence of both L-$A_{155}$ and L-$A_{10b}$, as well as each scrambled input individually, each correct input individually, both scrambled inputs together and both correct inputs together. An all D-DNA version of the AND gate showed substantial fluorescence regardless of input, as well as in the absence of any input, likely due to the degradation of the reporter (data not shown). The L-AND gate remained stable over the course of the experiment and was activated only in the presence of both correct inputs.

The L-DNA AND gate functions as a reporter for the detection of L-$OUT_{155}$ AND L-$OUT_{10b}$. To the best of our knowledge, this is the first example of an L-DNA system computing Boolean logic in live cells. Interestingly, the L-AND gate functions as well as, if not better, than either of the single component systems in live cells. It has low background, and ~3-fold activation compared to the 'no input' control. This could be due to experimental factors such as passage number or confluency, but it could also be due to the increasing complexity of the system. With either of the single component heterochiral strand displacement systems (miR-155 and miR-10b) any input agnostic leak can directly activate the downstream reporter

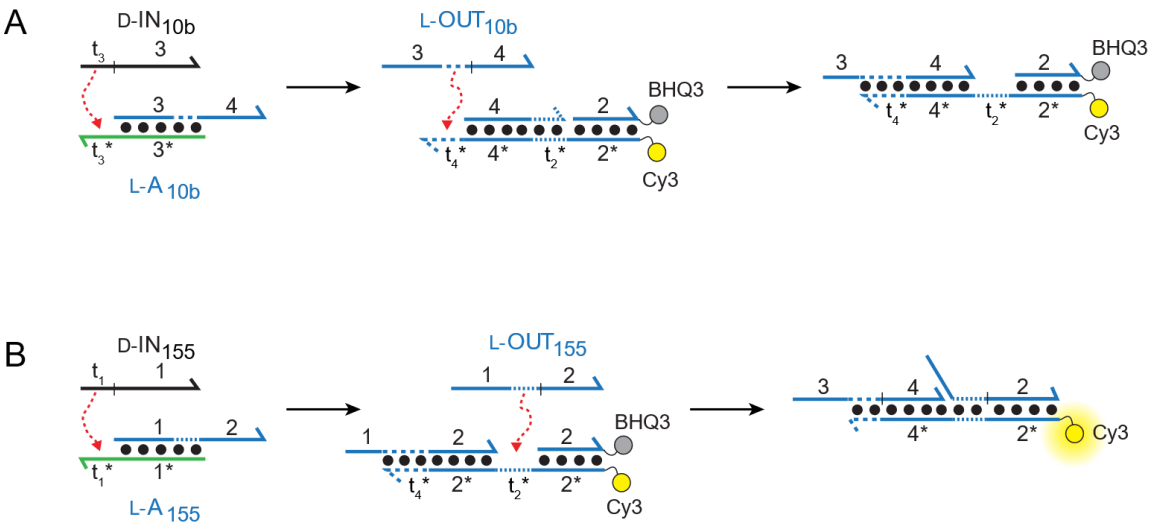



**Figure III-8. Scheme of L-AND gate activation.** (A) The first step of AND gate activation after the displacement of L-OUT$_{10b}$ from L-A$_{10b}$. (B) Second step of AND gate activation after displacement of L-OUT$_{155}$ from L-A$_{155}$.

| | Experimental Wells | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **DNA strand** | **+** | **-** | **No input (1)** | **10b scramble (2)** | **10b input (3)** | **155 scramble (4)** | **155 input (5)** | **Both scramble (6)** | **Both inputs (7)** |
| D or L-$R_{AND}$ | + | + | + | + | + | + | + | + | + |
| D or L-$OUT_X$ | + | - | - | - | - | - | - | - | - |
| D or L-$A_{155}$ and $A_{10b}$ | - | - | + | + | + | + | + | + | + |
| D or L-10b scramble | - | - | - | + | - | - | - | + | - |
| D or L-10b input | - | - | - | - | + | - | - | - | + |
| D or L-155 scramble | - | - | - | - | - | + | - | + | - |
| D or L-155 input | - | - | - | - | - | - | + | - | + |

**Table III-3.** Combinations of components mixed for experiments with the miR-155 and miR-10b L-AND gate.

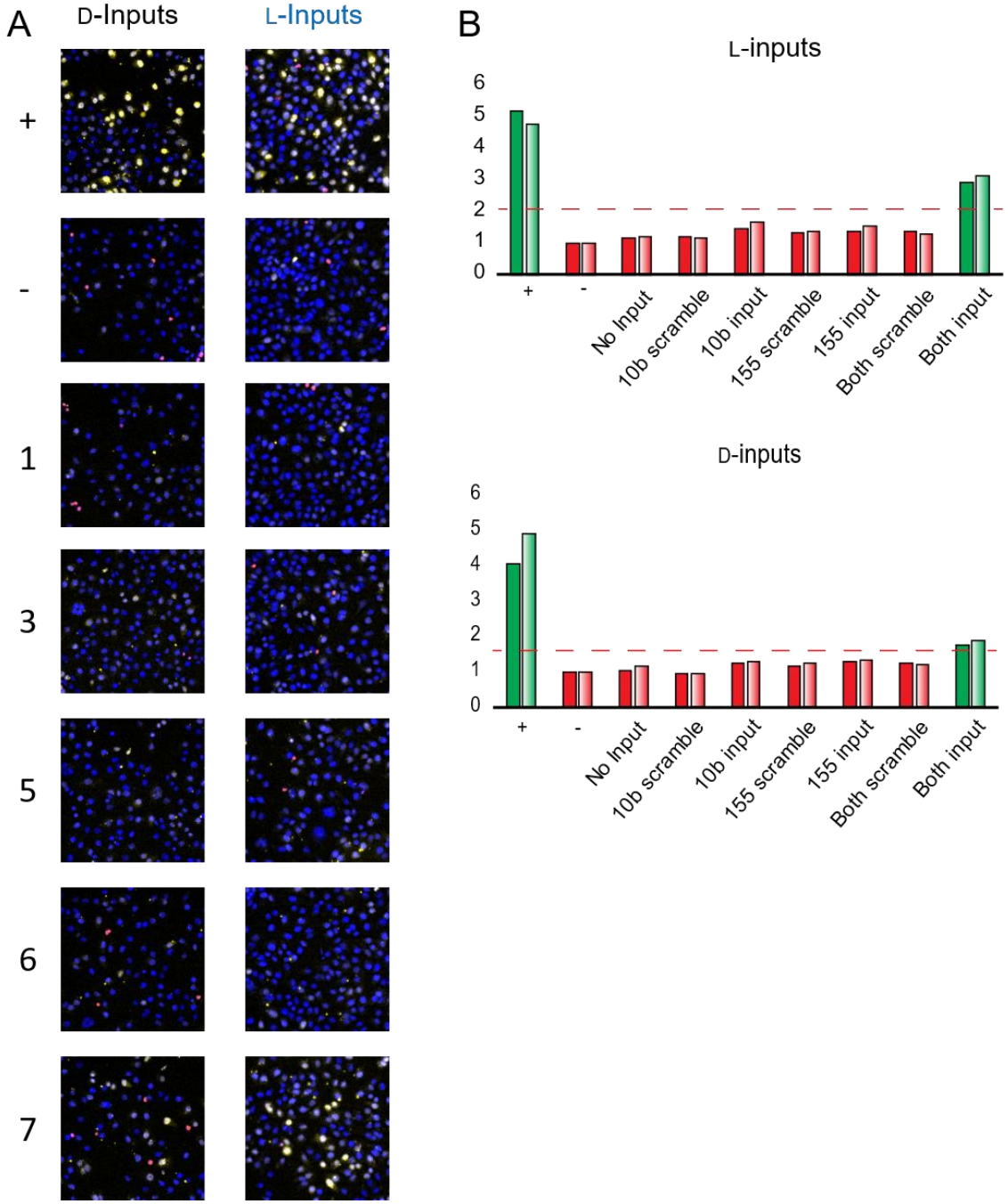



**Figure III-9. Activation of full L-AND system in live cells.** (A) Representative images of cells transfected with the indicated D or L-inputs. Labels correspond to mixtures indicated in Table III-3. (B) Quantification of the activation of each circuit in live cells. In each case, solid bars and gradient bars indicate biological duplicates. Bars were generated by normalizing the mean fluorescence of each population to a quenched L-R$_{AND}$ sample run on the same day.

duplexes. Whereas, in the multicomponent system, leak can only occur after non-specific interaction between L-$A_{155}$ AND L-$A_{10b}$ AND the reporter.

### 3.2 Conclusions

L-DNA strand displacement systems show high stability and semi-robust activation with respect to their D-DNA counterparts. Indeed, although the 2′-OMe-RNA system remained similarly stable in the cellular environment, the L-DNA heterochiral strand displacement displayed substantially higher activation both *in vitro* and *in vivo*. This highlights one of the most valuable design parameters for the generation of biostable DNA circuits: since only one step in heterochiral strand displacement is 'non-natural' (i.e. displacement of the output from the PNA/L-DNA heteroduplex), our near-quantitative understanding of DNA thermodynamics and hybridization can be directly applied to the generation of downstream sensor components. The data in this chapter demonstrates the comparative 'plug-and-play' design of two single component heterochiral strand displacement systems and a multicomponent heterochiral strand displacement system demonstrating Boolean AND logic.

Overall, the data presented in this chapter reveals some insights towards the design of L-DNA based computation devices, and highlights some of the challenges facing further adoption of this technology. L-DNA components still face many of the same challenges as their D-DNA counterparts, particularly regarding cellular entry and localization. While the L-DNA components used in these experiments were amenable to general transfection-reagent based delivery, different transfection reagent lots were inconsistent in terms of the

amount of material delivered and the overall activation. Some lots tended to have high fluorescence the nucleus, while many of the potential mRNA and microRNA targets reside predominately in the cytoplasm. Future work will build on the observations I've described in this chapter in two main ways. 1) A duplexes should be rigorously tested with potential reporters in live cells rather than just *in vitro* to help decrease leak in the absence of input, and 2) alternative transfection approaches should be strongly considered, in an effort to insure that most cells get receive a similar amount of each desired component.

## 3.3 Materials and Methods

### *3.3.1 DNA Design, synthesis and purification*

DNAs and RNAs were either purchased from Integrated DNA Technologies or prepared by solid-phase synthesis on an Expedite 8909 DNA/RNA synthesizer. DNA synthesis reagents, fluorophore phosphoramidites and D-nucleoside phosphoramidites were purchased from Glen Research, and L-nucleoside phosphoramidites were purchased from ChemGenes. Black Hole Quencher 2 (BHQ2) and Black Hole Quencher 3 (BHQ3) CPG resins were purchased from LGC Biosearch technologies. NHS ester version of the fluorescent dyes were purchased from Lumiprobe Life Science Solutions. PNAs were purchased from either PNA Bio Inc. or Panagene.

### *3.3.2 In vitro inhibition of heterochiral strand displacement by L2000*

Strand displacement reactions were evaluated in OptiMEM media, without FBS, mimicking transfection conditions with and without the addition of L2000. For the single

component heterochiral strand displacement systems (section 3.1.3, miR-10b and miR-155 systems), samples were prepared in the following fashion: D or L-$IN_{155}$ were separately diluted to 300 nM in 20 µL of pre-warmed OptiMEM, L-$A_{155}$ was diluted to 300 nM in 100 µL pre-warmed OptiMEM, and L-$R_{155.2}$ was diluted to 150 nM in 140 µL of pre-warmed OptiMEM. Samples were mixed by vortexing, then split in half. One half was diluted with an equal volume of OptiMEM. The other half was diluted with an equal volume of OptiMEM containing 0.2 µL L2000 (inputs), 1.0 µL L2000 (L-$A_{155}$), or 1.4 µL L2000 (L-$R_{155.2}$). DNAs were incubated with L2000 for 20 minutes at room temperature to form lipoplexes, then 10 µL of each complex was mixed according to table III-1 and diluted to a final volume of 30 µL as needed. The components without L2000 were mixed in the same fashion, and all samples were transferred to a 384-well black walled plate and analyzed by spectrofluorimetry. Signal generation was monitored over three hours to determine how long L2000 would prevent the circuit complexes from interacting (data shown in Figure III-3, L-155). The experiment with the miR-10b system was performed in the same fashion using the 10b components (see table A-3 for sequences).

For analysis of the multi-component heterochiral strand displacement system, an L-AND gate based on both the miR-155 and miR-10b systems, all steps were performed as above with minor adjustments: each input was diluted to 300 nM in 20 µL OptiMEM, both L-$A_{155}$ and L-$A_{10b}$ were diluted to 300 nM together in 140 µL OptiMEM, and L-$R_{AND}$ was diluted to 150 nM in 180 µL OptiMEM. Samples were vortexed, split in half, and diluted with either an equal volume of optiMEM alone or containing 0.4 µL L2000 (inputs), 2.1 µL L2000 (L-A mixture), or 3.6 µL L2000 (L-$A_{AND}$) and mixed as described

in table III-3. Samples were monitored by spectrofluorimetry as described above (Figure III-3, L-AND).

### *3.3.3 Cell culture*

HeLa (ATCC CCL-2) cells were grown in Dulbecco's Modified Eagle Medium (DMEM, Invitrogen) supplemented with 10% fetal bovine serum (FBS, Invitrogen), 25 mM HEPES (Invitrogen), 1X penicillin/streptomycin (Pen/Strep, Invitrogen) and incubated at 37 °C in a humidified atmosphere containing 5% $CO_2$. Cells were grown in cell culture treated T-25 flasks (Greiner Bio-One) and passaged as needed by gentle washing with 1 mL Dulbecco's Phosphate Buffered Saline (D-PBS) followed by detachment with 500 µL 0.25% Trypsin/EDTA (Invitrogen) for 5-10 minutes at 37 ºC. After incubation the trypsin/EDTA solution was inactivated by the addition of 2.5 mL of DMEM and a new flask containing 5 mL of fresh DMEM was seeded with either 500 or 250 µL cell suspension from the old flask (a 1:6 or 1:12 subculture ratio, respectively).

### *3.3.4 Transfections*

12-16 hours prior to transfection, cells were subcultured as described above, trypsinized, an aliquot was stained with 1 volume of either erythrosine B (ATCC) or trypan blue (VWR) and counted using a hemocytometer. Cells were then diluted to approximately 70,000 cells/mL in fresh DMEM without antibiotic, mixed gently by inversion, and 100 µL was added to each well of a 96-well plate (Greiner Bio-One, ~7,000 cells/well). The plates were tilted gently to distribute the cells evenly over the surface, left

in the laminar flow hood for 10-20 minutes to settle, then grown overnight at 37 ºC in a humidified atmosphere containing 5% $CO_2$. Transfections were performed once the cells reached 60-70% confluency the next day. Reproducibility of each experiment described in this chapter was assessed by performing all experiments in biological duplicate (e.g. at different passage number) on the same day.

For transfection of the single component heterochiral strand displacement system, circuit components that had first been validated *in vitro* were used in every experiment. Single component heterochiral strand displacement samples were prepared in the following fashion: D or L-$IN_{155}$ were separately diluted to 300 nM in 20 µL of pre-warmed OptiMEM, L-$A_{155}$ was diluted to 300 nM in 100 µL pre-warmed OptiMEM, and L-$R_{155.2}$ was diluted to 150 nM in 140 µL of pre-warmed OptiMEM. Samples were mixed by vortexing, then diluted with an equal volume of OptiMEM containing 0.4 µL L2000 (inputs), 2.0 µL L2000 (L-$A_{155}$), or 2.8 µL L2000 (L-$R_{155.2}$). DNAs were incubated with L2000 for 20 minutes at room temperature to form lipoplexes, then 40 µL of each complex was mixed according to table III-1 and diluted with OptiMEM to a final volume of 120 µL as needed. 60 µL of this mixture was added to HeLa cells at two different passage numbers (biological duplicate) and incubated for 1.5 hours. Each well was then washed, and the media was replaced with fresh DMEM (10% FBS, no antibiotic) and incubated for an additional 4.5 hours before imaging as described below. The experiment with the miR-10b system was performed in the same fashion using the 10b components (see table A-3 for sequences).

For the multi-component heterochiral strand displacement system, an L-AND gate based on both the miR-155 and miR-10b systems, all steps were performed as above with minor adjustments: each input was diluted to 300 nM in 20 µL OptiMEM, both L-$A_{155}$ and L-$A_{10b}$ were diluted to 300 nM together in 140 µL OptiMEM, and L-$R_{AND}$ was diluted to 150 nM in 180 µL OptiMEM. Samples were vortexed then diluted with an equal volume of OptiMEM containing 0.8 µL L2000 (inputs), 5.4 µL L2000 (L-A mixture), or 7.2 µL L2000 (L-$A_{AND}$) and mixed as described in table III-3. Samples were diluted to a final volume of 120 µL with OptiMEM and transfected as described above.

### *3.3.5 Quantitative determination of fluorescence by flow cytometry*

For fluorescence measurement using the flow cytometer, cells were trypsinized after microscopy and resuspended with OptiMEM to a final volume of 230 µL. Cells were then analyzed using a BD Accuri C6 Plus flow cytometer equipped with an FL2 – PE-H filter (Ex: 488 nm/Em: 533/40 nm). All data was collected at a flow rate of 66 µL/min and in every case 200 µL was sampled. The geometric mean of the FL2 signal for each experiment was determined using the Flowjo software. Mean fluorescent intensity was normalized to the L-R only sample and plotted as a bar graph.

### *3.3.6 Cell viability*

In order to determine cells that had compromised plasma membranes, all cells were treated with DRAQ7 and Hoechst before imaging. DRAQ7 is cell impermeable and stains the DNA of cells with compromised plasma membranes. The Hoechst dye is cell

permeable and stains the DNA of all cells. Images were acquired using the blue and red filters (DAPI and Cy5 light cubes, respectively) and cells were counted using Celleste (Invitrogen) to determine the total cells (blue) and the number of dead cells (red). Cytotoxicity was determined from the ratio of red cells/total number of cells.

### 3.2 References


1. Adleman, L. M., Molecular computation of solutions to combinatorial problems. *Science* **1994,** *266* (5187), 1021.

2. Seelig, G., et al., Enzyme-Free Nucleic Acid Logic Circuits. *Science* **2006,** *314* (5805), 1585.

3. Qian, L.; Winfree, E., Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **2011,** *332* (6034), 1196.

4. Srinivas, N., et al., Enzyme-free nucleic acid dynamical systems. *Science* **2017,** *358* (6369), eaal2052.

5. Groves, B., et al., Computing in mammalian cells with nucleic acid strand exchange. *Nat. Nanotechnol.* **2015,** *11*, 287-295.

6. Chen, Y.-J., et al., DNA nanotechnology from the test tube to the cell. *Nat. Nanotechnol.* **2015,** *10*, 748.

7. Wlotzka, B., et al., In vivo properties of an anti-GnRH Spiegelmer: An example of an oligonucleotide-based therapeutic substance class. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99* (13), 8898.

8. Vater, A.; Klussmann, S., Turning mirror-image oligonucleotides into drugs: the evolution of Spiegelmer therapeutics. *Drug Discovery Today* **2015,** *20* (1), 147-55.

9. Xu, Y.; Szoka, F. C., Mechanism of DNA Release from Cationic Liposome/DNA Complexes Used in Cell Transfection. *Biochemistry* **1996,** *35* (18), 5616-5623.

10. Cardarelli, F., et al., The intracellular trafficking mechanism of Lipofectamine-based transfection reagents and its implication for gene delivery. *Sci. Rep.* **2016,** *6*, 25879.

11. Lacroix, A., et al., Uptake and Fate of Fluorescently Labeled DNA Nanostructures in Cellular Environments: A Cautionary Tale. *ACS Cent. Sci.* **2019,** *5* (5), 882-891.

12. Groves, B., et al., Computing in mammalian cells with nucleic acid strand exchange. *Nat. Nanotechnol.* **2015,** *11*, 287.

13. Afonin, K. A., et al., Activation of different split functionalities on re-association of RNA–DNA hybrids. *Nat. Nanotechnol.* **2013,** *8*, 296.

14. Zadeh, J. N., et al., NUPACK: Analysis and design of nucleic acid systems. *J. Comput. Chem.* **2011,** *32* (1), 170-173.

# CHAPTER IV

# CHARACTARIZING THE KINETICS OF HETEROCHIRAL STRAND DISPLACEMENT REACTIONS

One of the most valuable attributes of DNA strand displacement systems is our near quantitative understanding of the thermodynamics and kinetics of DNA hybridization. This makes WC base pairing a powerful, programmable interaction modality that allows researchers to develop incredibly complex systems, from static nanoscale structures[1] to multi-layered computational circuits[2-5]. Our heterochiral strand displacement system utilizing a PNA/DNA heteroduplex is unique among biostable DNA devices in that there is only one step that is "unnatural." With other nucleic acid modifications such as phosphorothioate linkages, 2′-OMe RNA, or locked nucleic acids (LNAs) the thermodynamic and kinetic properties of hybridization are dramatically affected compared to unmodified DNA[6-8]. Such modifications must be present within every DNA strand in a given system for these devices to function in a biological environment, therefore they ultimately undermine the established simplicity of DNA strand displacement. However, an L-DNA system will function identically to a D-DNA system of the same sequence and as such it can be designed rationally or with computational tools in the same manner as systems of the natural chirality. In this fashion, it is only necessary to experimentally optimize the heterochiral strand displacement event itself. I established in chapter II that the reaction rate with a heterochiral input is substantially slower than with a homochiral input[9], and any strand displacement reaction involving a PNA toehold is slower than the established rates for DNA strand

displacement[10-11]. In this chapter I systematically tested a series of inputs that will enable a better understanding of the difference in rate between strand displacements utilizing a homochiral vs a heterochiral input. This will help inform the design of future heterochiral strand displacement devices. First, I tested inputs with various toehold lengths to identify when, if ever, the rate of heterochiral strand displacement reaches parity with homochiral strand displacement. Second, I have chosen a series of mismatches to test the tolerance of heterochiral strand displacement for incorrect base pairing of an input strand within the toehold or branch migration domains. These inputs are especially informative because many disease related RNAs within the cell may only differ from the natural target by a single nucleotide[12], and the sensitivity of heterochiral strand displacement – based biosensors to such changes is an important consideration. Finally, I have tested a series of truncated input strands that can't fully invade the branch migration domain. Inputs designed in this manner mimic the DNA circuit design known as toehold exchange[4, 10], enabling the potential development of catalytic systems or modular reporters without sequence constraints incurred due to the necessity of target complementarity. Most of these inputs were tested with two different length branch migration domains, one containing 20 base pairs and another with 16 base pairs, although some of the data with the short branch migration domain is still being finalized.

After spending so much time with these various heterochiral strand displacement systems, I wanted to identify where the rate penalty between the homochiral and heterochiral strand displacement system originates, as this might translate directly to the improved design and performance of future systems. There are multiple steps within the

strand displacement reaction that impact the overall kinetics of strand displacement. All of these steps can be conceptually simplified to four distinct timescales: toehold nucleation and zippering, fraying, initiation of branch migration, and the overall process of branch migration[13] (Figure IV-1). The goal of my work in this chapter is to begin the optimization of next generation heterochiral strand displacement devices through a better understanding of PNA enabled heterochiral strand displacement, and to identify the kinetic implications of chiral inversion during heterochiral strand displacement. These studies build on the design of the novel thresholding reaction described in section 2.1.2, whereby the kinetic penalty of chiral inversion was used to ensure that the heterochiral strand displacement reaction only occurred in the presence of a large excess of input strand.

The strand-displacement systems previously discussed were designed with 6 base toeholds and sequences corresponding potential endogenous targets (miR-155, miR-10b and the MnSOD mRNA), and those systems were diverse enough to infer some sequence-specific effects on heterochiral strand displacement. In the current chapter we utilized a system based on sequences reported by Zhang and Winfree for the kinetic analysis of DNA strand displacement reactions[10]. The sequences were chosen to have no secondary structure in single stranded regions in an effort to decouple the approximately second order kinetics of strand displacement from the first order kinetics of secondary structure unfolding. An important design consideration for this effort is minimizing the number of G bases in single stranded regions due to the strength of the G-C base pair, the stability of G-T wobble pairs, and the potential for G-quadruplex formation. The PNA strands in this system both contained a 10 base toehold and all inputs, $A_{kinetics}$ and $R_{kinetics}$ were gel

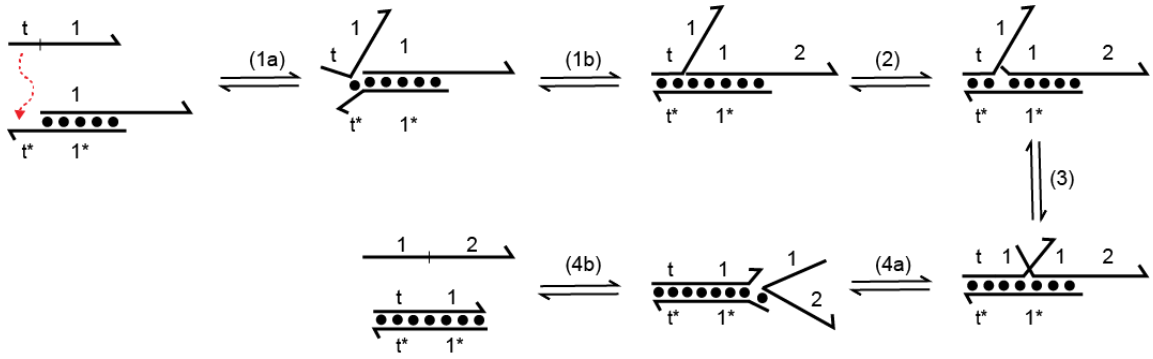



**Figure IV-1. Individual steps of DNA strand displacement.** Up to this point the overall process of strand displacement has been discussed in general terms. For this section, it will be useful to consider the individual steps of strand displacement in a little more detail. It is important to note that each of these steps represent interactions in equilibrium. Strand displacement is initiated by toehold nucleation (1a), where the input strand forms 1-3 base pairs with the available toehold. This transient interaction can either dissociate (reverse reaction) or "zip up" and form the fully hybridized toehold (1b). Fraying at the ends of DNA duplexes (2) frees an additional base at the branch migration junction that allows the input strand to bind. In this way the input strand has replaced one base pair of the original branch migration domain. Fraying and individual steps of branch migration can occur in either direction (3) leading to a random walk of the input strand back and forth through the branch migration domain. Once the input reaches the end of the branch migration domain (4a), the final base can dissociate, and branch migration is completed (4b). Due to fraying at the ends of DNA duplexes, step 4b is indeed reversible but at a rate so slow that it is often considered negligible.

purified to ensure optimal circuit performance. After purification, concentrations of the $A_{kinetics}$ and $R_{kinetics}$ stereoisomers were approximated by UV absorbance then experimentally normalized to the fluorescence of the unquenched fluorophore labeled strand. For the remainder of this chapter these duplexes will be referred to as D or L-$A_{kinetics}$ and D or L-$R_{kinetics}$. Sections 4.1.1 and 4.1.2 will discuss differences between two versions of the $A_{kinetics}$ duplex, one with a long branch migration domain (20 base pairs) and one with a shorter branch migration domain (16 base pairs). These duplexes will be referred to as D or L-A_$L_{kinetics}$ and D or L-A_$S_{kinetics}$, respectively. and they were prepared using either the full length PNA or one truncated by four base pairs (Figure IV-2).

## 4.1 Results

### *4.1.1 The effect of toehold length on the kinetics of heterochiral strand displacement*

Toehold length plays an important role in the design of DNA strand displacement systems and can modulate the rate of such reactions by over 6 orders of magnitude in traditional all DNA systems[10-11]. Based on the decreased rate observed in systems containing a PNA toehold, especially with a heterochiral input, it was expected that by increasing the toehold length we could accelerate heterochiral strand displacement reactions *in vitro* and in live cells. Toehold lengths were modulated by truncating the inputs such that they can only interact with 2, 4, 6, 8 or 10 bases of the PNA toehold (Figure IV-2). Strand displacement reactions containing all-DNA components experience a rate enhancement of several orders of magnitude (~$10^{6.8}$) between toeholds containing 2 to >10 base pairs[10, 13]. While this enhancement is not fully explained by the expected

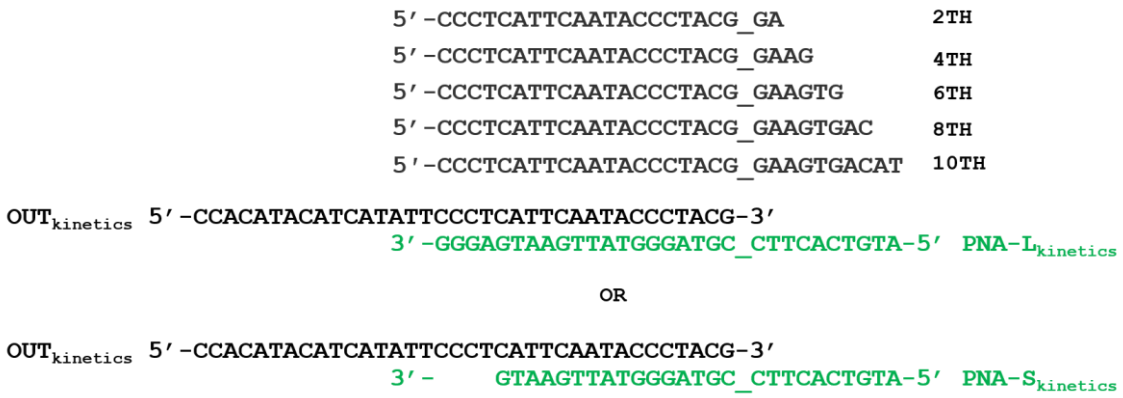
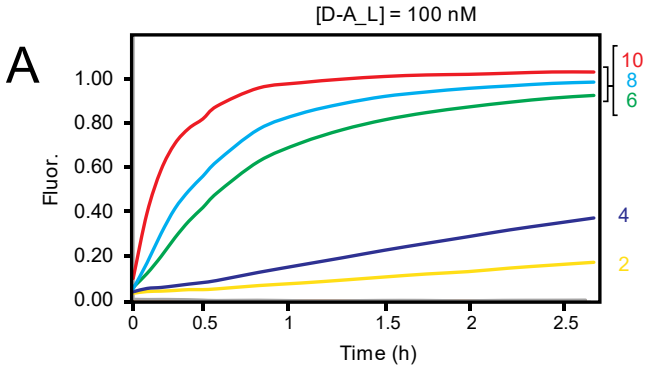


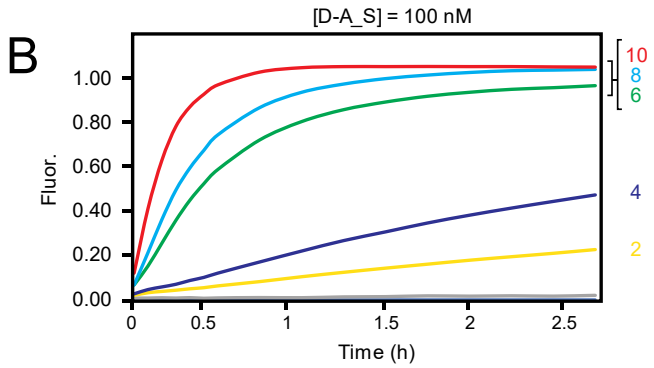

**Figure IV-2. Sequences and names of the toehold inputs discussed in this section.** Each of these inputs was tested for its ability to displace $OUT_{kinetics}$ from the PNA substrate strand. Two different PNAs were tested in these experiments, PNA_$L_{kinetics}$ with a 20 nucleotide (nt) branch migration domain and PNA_$S_{kinetics}$ with a 16 nt branch migration domain. When PNA_$L_{kinetics}$ or PNA_$S_{kinetics}$ are hybridized to D or L-$OUT_{kinetics}$, the resulting duplex is referred to as D or L-A_$L_{kinetics}$ or D or L-A_$S_{kinetics}$, respectively.

hybridization rates of DNA at various toehold lengths[10, 13], in the context of strand displacement reactions it is nevertheless a valuable tool for the design of DNA nanodevices and sensors. This rate enhancement plateaus for traditional DNA systems at approximately a 6-base toehold at 23 ºC[10], but we expected that heterochiral strand displacement systems might benefit from longer toeholds due to the increased reaction temperature (37 ºC) and from previous data suggesting a reduced rate of PNA/DNA strand displacement[9]. Reactions were prepared as described in section 4.3.5. Briefly, reactions were mixed in a 30 µL final volume with 150 nM of the D-input, 100 nM of the D or L-$A_{kinetics}$ duplex, and 300 nM of the D or L-$R_{kinetics}$ duplex with 300 mM NaCl, 1 mM EDTA and 50 mM tris pH 7.6.

For the homochiral system, calculated rate constants were strongly dependent on toehold length for both the long and short branch migration domains. In each case the 6, 8 and 10 base toeholds approximately reached the maximum possible signal, with the 2 and 4 base toeholds being substantially slower (Figure IV-3 A and B). The rate dependence on toehold length is roughly exponential, as has been previously described for all-DNA strand displacement reactions[10] (Figure IV-3 C). Branch migration domain length seems to play a minor role in the overall reaction rate, with the shorter branch migration domain being marginally faster within each toehold comparison. In contrast, the heterochiral reactions only reached completion for the 8 and 10 base toeholds, the 6 base toehold was slower than the 4 base toehold in the corresponding homochiral system, while the heterochiral inputs with 2 or 4 base toeholds were barely active (Figure IV-4 A and B). Once again, the reaction system with the shorter branch migration domain was slightly




| Toehold Length | k ($M^{-1} s^{-1}$) |
|---|---|
| 10 | 9624 |
| 8 | 4352 |
| 6 | 2150 |
| 4 | 695 |
| 2 | 275 |




| Toehold Length | k ($M^{-1} s^{-1}$) |
|---|---|
| 10 | 11562 |
| 8 | 4929 |
| 6 | 3020 |
| 4 | 927 |
| 2 | 423 |

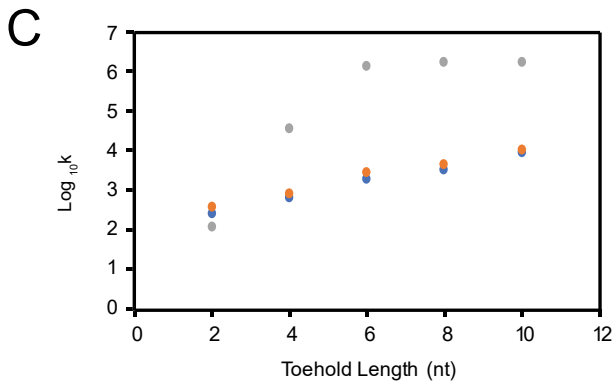
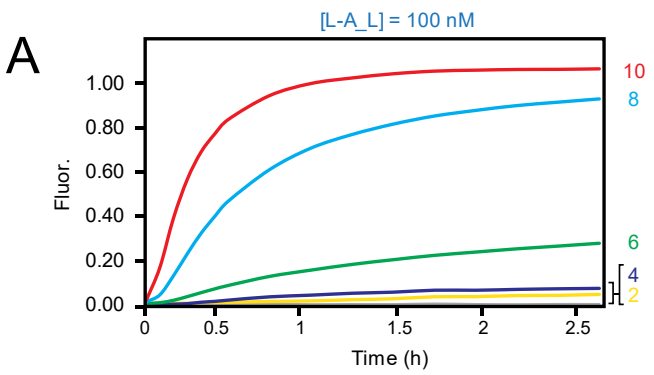
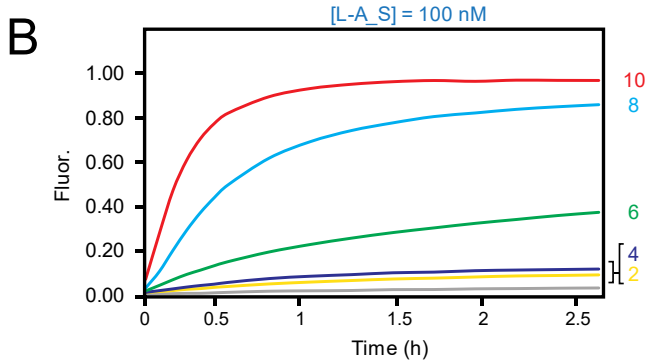



**Figure IV-3. Dependence of homochiral strand displacement on toehold length.** (A) D-A_L$_{kinetics}$ displaced by inputs of various toehold lengths. Toehold length in nt is shown on the right of the graph. (A) D-A_S$_{kinetics}$ displaced by inputs of various toehold lengths. Toehold length in nt is shown on the right of the graph. (C) Semilogarithmic plot of toehold length vs log(k). Blue dots represent rates extracted from plot (A), orange dots represent rates pulled from plot (B), and grey dots represent the all-DNA version of this system reported by Zhang and Winfree in ref. 10. Adapted with permission from the American Chemical Society.




| Toehold Length | k ($M^{-1}$ $s^{-1}$) |
|---|---|
| 10 | 6097 |
| 8 | 2268 |
| 6 | 578 |
| 4 | 159 |
| 2 | 85 |




| Toehold Length | k ($M^{-1}$ $s^{-1}$) |
|---|---|
| 10 | 7024 |
| 8 | 4228 |
| 6 | 1534 |
| 4 | 217 |
| 2 | 175 |

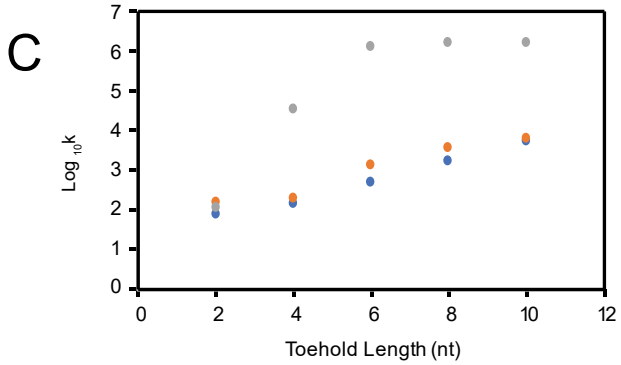



**Figure IV-4. Dependence of heterochiral strand displacement on toehold length.** (A) L-A_$L_{kinetics}$ displaced by inputs of various toehold lengths. Toehold length in nt is shown on the right of the graph. (A) L-A_$S_{kinetics}$ displaced by inputs of various toehold lengths. Toehold length in nt is shown on the right of the graph. (C) Semilogarithmic plot of toehold length vs log(k). Blue dots represent rates extracted from plot (A), orange dots represent rates pulled from plot (B), and grey dots represent the all-DNA version of this system reported by Zhang and Winfree in ref. 10. Adapted with permission from the American Chemical Society.

faster than the longer branch migration domain with the same inputs. Interestingly, the rate constants for the heterochiral system display a similarly exponential dependence on toehold length (Figure IV-4 C). It would be interesting to follow up these experiments with even shorter branch migration domains to more accurately determine the dependence of reaction rate on branch migration length, especially considering the shortest test branch migration domain we've tested previously is 13 base pairs (miR-10b system, chapter 2).

One of the most surprising observations made during these experiments is that the homochiral reactions with a 6 base toehold were substantially slower than every other system PNA/DNA system we've previously designed (rate constants 4-7-fold smaller than those I reported in chapter II). This goes against what we initially expected, since the toehold in this experiment contains 3 strong base pairs (50% G – C content) which is the same as the miR-10b system (the fastest of the 3 systems discussed in chapter 2, Figure IV-5). We expect there are some sequence-specific justifications that might explain this disparity: while the length of the branch migration domain doesn't seem to have a large impact on the rate of strand displacement, it is certainly likely that base pair composition does. Initiation of branch migration has been shown to be slow due to the steric impact of generating a second single stranded overhang[13], and that is possibly compounded in this system since the invading strand is impeded by two strong base pairs immediately at the branch migration junction. However, while the heterochiral rates show a similar trend they seem less impacted by whatever is causing this discrepancy, with rate constants only 1.1-5-fold slower than those reported in chapter 2. This could be tested by purchasing a new PNA with a branch migration domain containing similar GC content but shifted away

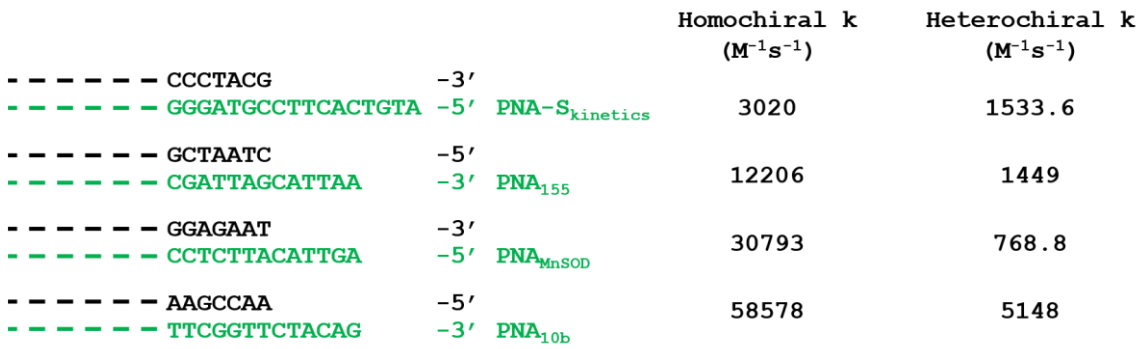



**Figure IV-5. Potential effects of sequence on strand displacement rate.** It is important to note that D-A_$S_{kinetics}$ and D-$A_{MnSOD}$ have 5′ toeholds as opposed to the 3′ toeholds found in D-$A_{155}$ and D-$A_{10b}$. It is not clear at this time what effect toehold polarity might have on strand displacements with a PNA toehold, but it will likely have some effect. The clearest trend within the homochiral rate constants is that the placement of strong base pairs at the branch migration domain is inversely proportional to the rate (more G – C base pairs near the junction predicts slower rates).

from the initial junction of the branch migration domain to see what impact these changes have on the reaction rate.

### *4.1.2 The effect of mismatch position on the kinetics of heterochiral strand displacement*

The effect of mismatched base pairs within an input strand participating in toehold mediated strand displacement has been studied with all-DNA strand displacement reactions over the last decade[14-15]. Mismatches at various positions within the branch migration domain or the toehold domain can affect the rate of strand displacement reactions by multiple orders of magnitude[14]. This may provide impetus for the incorporation of specific mismatches to fine tune the rates of complex molecular computation systems or help inform the design of DNA strand displacement probes that can efficiently discriminate between single nucleotide polymorphisms (SNPs) in live cells. Since PNA has been reported to be more sensitive to mismatches than standard DNA probes[16-17], we hypothesized that strand displacement systems with a PNA substrate strand would show a more pronounced affect with mismatch-bearing input strands. Additionally, with the substantially reduced kinetics of heterochiral strand displacement reactions utilizing a PNA toehold we predicted this type of system could be a powerful tool for mismatch discrimination if designed appropriately. To test this, a series of inputs containing mismatches at single positions (and one input containing two mismatches within the toehold binding domain) were generated from the 8-base toehold invader described in the previous section (Figure IV-6). The effect of mismatch position on the

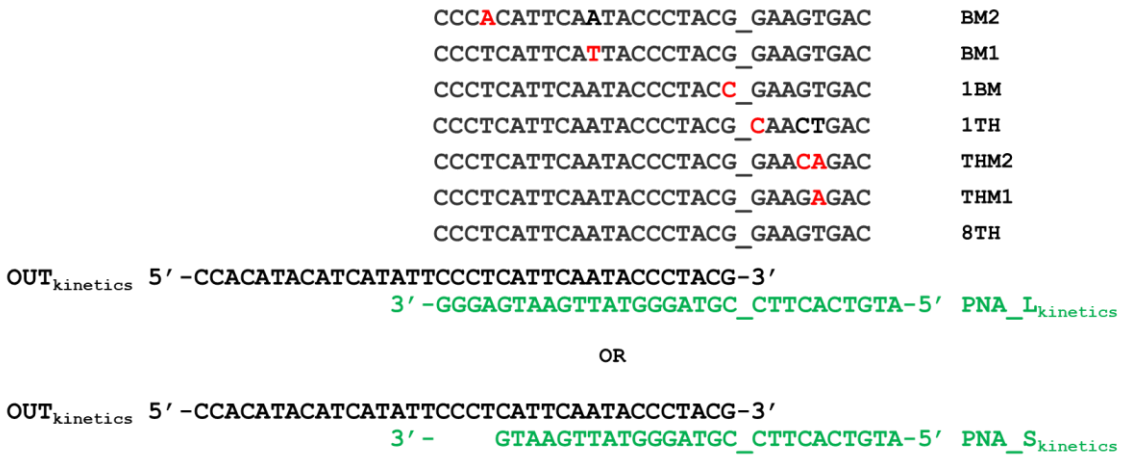



**Figure IV-6. Sequences and names of the mismatch inputs discussed in this section.** Each of these inputs was tested for its ability to displace $OUT_{kinetics}$ from the PNA substrate strand. Two different PNAs were tested in these experiments, $PNA\_L_{kinetics}$ with a 20 nucleotide (nt) branch migration domain and $PNA\_S_{kinetics}$ with a 16 nt branch migration domain. When $PNA\_L_{kinetics}$ or $PNA\_S_{kinetics}$ are hybridized to D or L-$OUT_{kinetics}$, the resulting duplex is referred to as D or L-$A\_L_{kinetics}$ or D or L-$A\_S_{kinetics}$, respectively. Bases highlighted in red indicate a mismatch within the resulting input-PNA duplex.

rate of strand displacement was systematically tested with both the homo- and heterochiral systems, as well as 16 and 20 base pair branch migration domains as described previously. Based on literature regarding PNA/DNA hybridization, and the effect of mismatches on all-DNA strand displacement systems, we had some *a priori* expectations on how these mismatches would behave in these systems. Mismatches within the branch migration domain (1BM, BM1 and BM2) were expected to more strongly decrease the reaction rate the closer they are to the toehold domain[14]. Mismatches in the toehold (1TH, TH1 and TH2) should behave in a similar fashion, toeholds closest to the incumbent duplex should show more prominent retardation of the reaction rate. Larger effects closer to the beginning of the branch migration domain likely arise from an understood thermodynamic penalty for initiating branch migration[13], arising from the formation of a new single stranded domain as the incumbent strand begins to be displaced (see Figure IV-1, step 3 for reference). Perturbations directly adjacent to the branch migration junction should have a more prominent impact on reaction kinetics.

Unsurprisingly, mismatches within the middle or near the end of the branch migration domain (BM1 and BM2) had little effect on the rate of strand displacement for the homochiral system, regardless of branch migration domain length (D-A_$L_{kinetics}$ and D-A_$S_{kinetics}$, Figure IV-7 A and B). This is consistent with literature investigating mismatch discrimination in all DNA systems[14], and suggests an alternative dissociation pathway for release of the output strand rather than the complete step-by-step displacement to the end of the branch migration domain (i.e. once the input strand has invaded enough of the branch migration domain, the incumbent strand may fall off due to melting of the

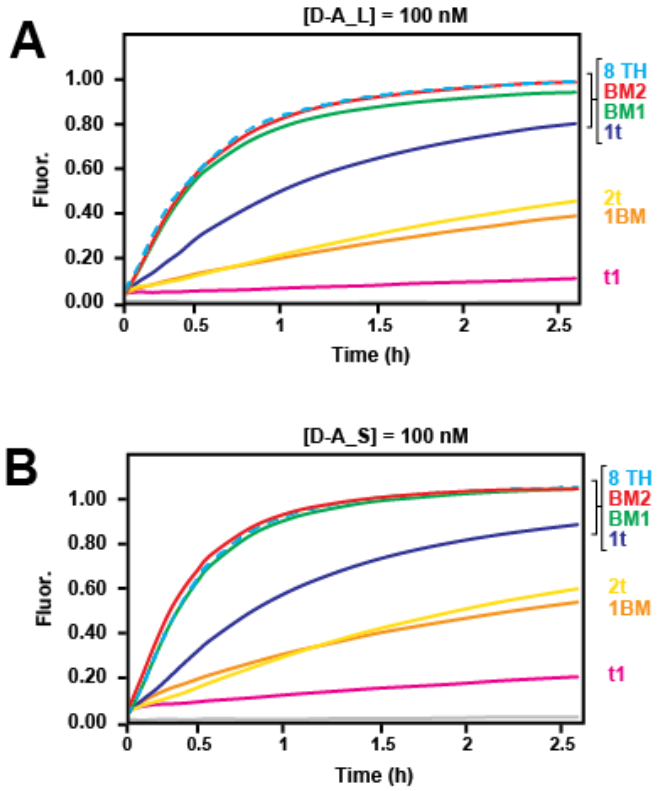



**Figure IV-7. The kinetics of homochiral strand displacement are affected by mismatches within the input strand.** In each graph the trace from the 8 TH input is shown for comparison (dashed blue line). (A) The ability of a mismatched input strand to undergo strand displacement through a long branch migration domain is dependent on the position of the mismatch. (B) The ability of a mismatched input strand to undergo strand displacement through a short branch migration domain is dependent on the position of the mismatch.

remaining base pairs rather than completion of the step by step process of branch migration). The TH1 mismatch, where the input contains a mismatch near the middle of the toehold, displays a substantial reduction in reaction rate for the homochiral system, likely due in part to the reduced toehold – input stability. The TH2 mismatch decreases the rate even further, as expected, by adding a second mismatch adjacent to the first. Mismatch 1BM effected the rate by a similar magnitude as TH2, even though the toehold domain for this invader is fully complementary to the target and there is only one mismatched site compared to two in the TH2 input. Early mismatches within the branch migration domain have been demonstrated to have a large effect on the kinetics of strand displacement[14], generally thought to occur due to the thermodynamic penalty of harboring a mismatch within the duplex. This further decreases the likelihood of successfully initiating branch migration, increasing the possibility of fraying and spontaneous dissociation of the toehold – invader interaction. Finally, the 1TH mismatch almost completely retards strand invasion. This is slightly surprising considering its similarity to the 1BM mismatch, both are directly adjacent to the branch migration initiation point and both are C – C mismatches. Computational studies in the literature suggest that coaxial stacking between base pairs between the invading duplex and the incumbent duplex play an important role in the mechanism of branch migration[13] as well as the kinetics of DNA hybridization[18], and it seems likely that a mismatch at this position would more strongly disrupt these interactions compared to the 1BM mismatch.

For the heterochiral system (L-A_$L_{kinetics}$ and L-A_$S_{kinetics}$, Figure IV- 8 A and B), a distal mismatch (BM2) within the long branch migration domain behaved very similarly

to the fully complementary 8 base toehold. Again, this is most likely due to the dissociation of the incumbent strand before the mismatch is reached. This same effect is seen for BM1 with the short branch migration domain, which is consistent with this reasoning. The BM1 mismatch is still relatively well tolerated with the longer branch migration domain, reaching ~80% of the max signal over the time course. While many of the other mismatches were partially tolerated (TH1, TH2 and 1BM) in the homochiral system all of them reached <20% of the maximum signal with the heterochiral system (Figure IV-8 A and B). This information may be invaluable in the design of L-DNA based probes for the discrimination of point mutations within a cellular context. Future work will focus in part on mismatch discrimination with RNA inputs, but the work in this section suggests that heterochiral strand displacement systems can be specifically designed to discriminate against relevant mismatches based on placement of the mismatch within the PNA strand.

### *4.1.3 The effect of invader length on the kinetics of heterochiral strand displacement*

While it seems logical that branch migration would end with the input strand completely displacing the incumbent strand, recent kinetics studies suggest that the final steps of branch migration are likely skipped through an alternative dissociation driven pathway[10, 14]. That is, once the invader displaces enough bases on the substrate strand, the remaining bases can spontaneously dissociate and release the incumbent strand. This observation was first described in the development of toehold exchange reactions, a subset of toehold mediated strand displacement systems in which the invader strand is shorter

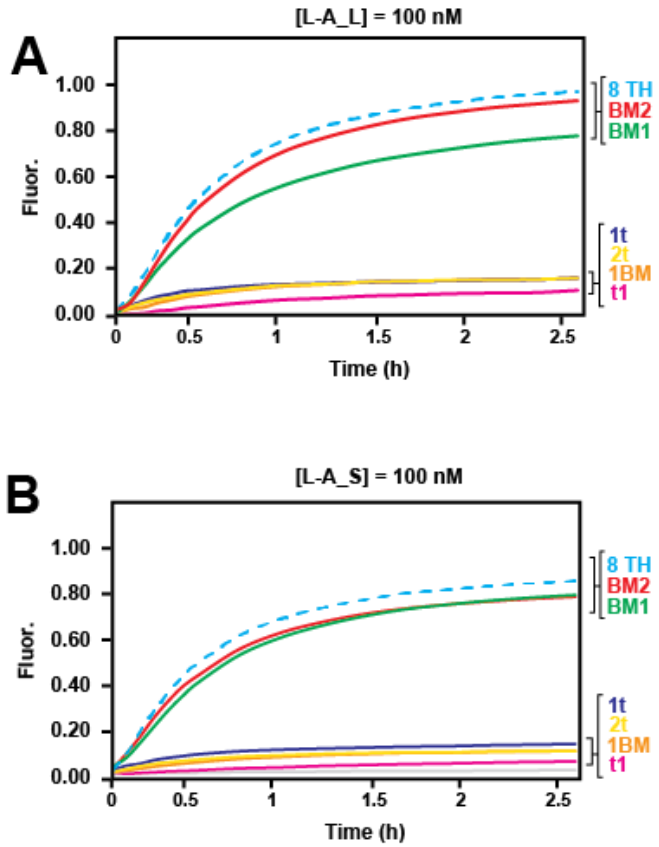



**Figure IV-8. The kinetics of heterochiral strand displacement are affected by mismatches within the input strand.** In each graph the trace from the 8 TH input is shown for comparison (dashed blue line). (A) Only mismatches at the distal positions of the longer branch migration domain are tolerated. (B) Only mismatches at the distal positions of the shorter branch migration domain are tolerated.

than the branch migration region of the substrate duplex[4-5, 10]. In these toehold exchange reactions the spontaneous dissociation of the incumbent strand after partial invasion by the invader reveals a new toehold on the substrate strand, and the length of each of these toeholds with respect to the other will lead to a back-and-forth equilibrium of strand displacement unless the system is perturbed by additional components (such as a downstream reporter duplex sequestering the OUT strand). Alternatively, the spontaneous dissociation of the incumbent strand may enable the design of reporter systems with complete sequence independence from the input strand. The ability of inputs of different lengths to effectively displace the incumbent strand was evaluated for both the homo- and heterochiral systems (Figure IV-9).

So far, the effects of progressively shortened input strands containing an 8-base toehold (-4, -6 and -8 bases compared to the full-length branch migration domain) have been evaluated with the 20 base pair branch migration domain. These reactions were compared to the full-length input with an 8-base toehold as a control. For the homochiral system, the results were mostly as expected (D-A_$L_{kinetics}$, Figure IV-10 A). The longest input (-4) was the fastest, the -6 input was substantially slower, and the -8 input was slow enough that it did not reach completion during the time course. However, both the -4 and -6 inputs reached a higher signal than should be possible with this experimental design, based on the D-A_$L_{kinetics}$ duplex being the limiting reagent. In these strand displacement reactions, the waste duplex generated is usually fully base paired and thus considered unlikely to participate any further in the reaction. Since the waste products of these reactions contain a new 4 or 6 nt single stranded region at the 3′ end of the PNA, it may


**Figure IV-9. Sequences and names of the short inputs discussed in this section.** Each of these inputs was tested for its ability to displace $OUT_{kinetics}$ from the long PNA substrate strand ($PNA\_L_{kinetics}$) with a 20 nt branch migration domain. When $PNA\_L_{kinetics}$ is hybridized to D or L-$OUT_{kinetics}$, the resulting duplex is referred to as D or L-$A\_L_{kinetics}$.

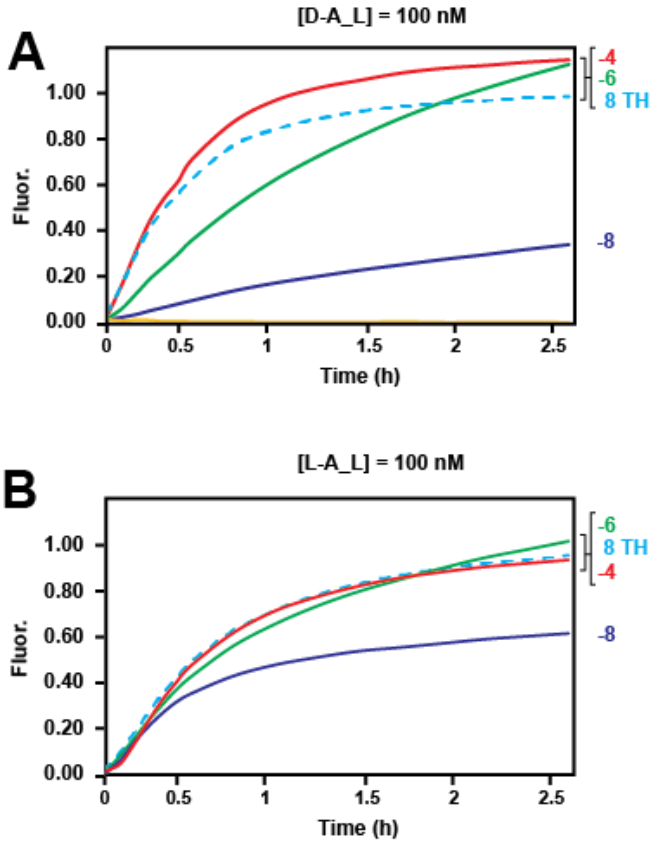



**Figure IV-10. The kinetics of strand displacement are affected by truncating the input strand.** In each graph the trace from the 8 TH input is shown for comparison (dashed blue line). (A) Homochiral reaction with D-A_L$_{kinetics}$. (B) Heterochiral reaction with L-A_L$_{kinetics}$.

be possible that they tend to leak more in the presence of excess reporter duplex making it possible to generate more fluorescence than otherwise expected. While this reaction pathway seems unlikely, this can be tested by annealing these truncated inputs to PNA_$L_{kinetics}$, which would be equivalent to the waste duplex created during the strand displacement reaction tested here. This duplex could be incubated with the D-$R_{kinetics}$ module to test for spurious fluorescence (Figure IV-11).

The results from the heterochiral system (L-A_$L_{kinetics}$, Figure IV-10 B) were surprising and have interesting implications for the design of future heterochiral strand displacement systems. The -4 input behaved as expected, almost exactly mirroring the reaction with the fully complementary input (8TH, see section 4.1.1) having the same length toehold (8 nt). However, both the -6 input and the -8 input have faster initial rates than the same inputs with the homochiral system. These are the only examples seen so far having this behavior, and further testing needs to be done to determine how general this phenomenon is. The reactions with these truncated inputs have the same initial rates as the heterochiral reaction with the fully complementary 8 nt toehold, and this observed rate acceleration may be due to changes in the equilibrium of the "post" branch migration intermediate. For the homochiral system, both the invader and the incumbent strand are of the same chirality. Naively, the truncated input displaces a portion of the branch migration domain until it reaches an intermediate state consisting of a fully hybridized invader and a partially hybridized incumbent strand (Figure IV-12 A). From this intermediate state, the incumbent strand can either dissociate (followed by activation of the reporter or re-nucleation onto the still-available toehold) or it can begin branch migration in the reverse

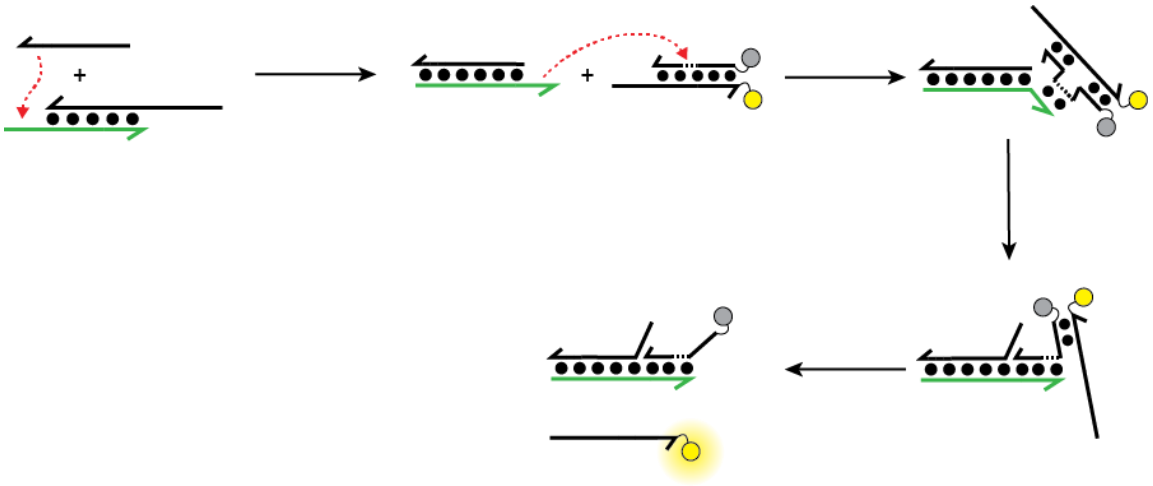


**Figure IV-91. Possible pathway for the generation of excess signal with truncated inputs in the homochiral displacement pathway.** Homochiral strand displacement with a truncated homochiral input. The revealed 3′ toehold on the PNA is complementary to a region within the quencher strand. Certain PNA sequences have been demonstrated to possess some strand invasion capability, so it may be possible that the higher than expected signal seen with many of the truncated inputs is due to alternate invasion pathways based on the remaining single stranded PNA.

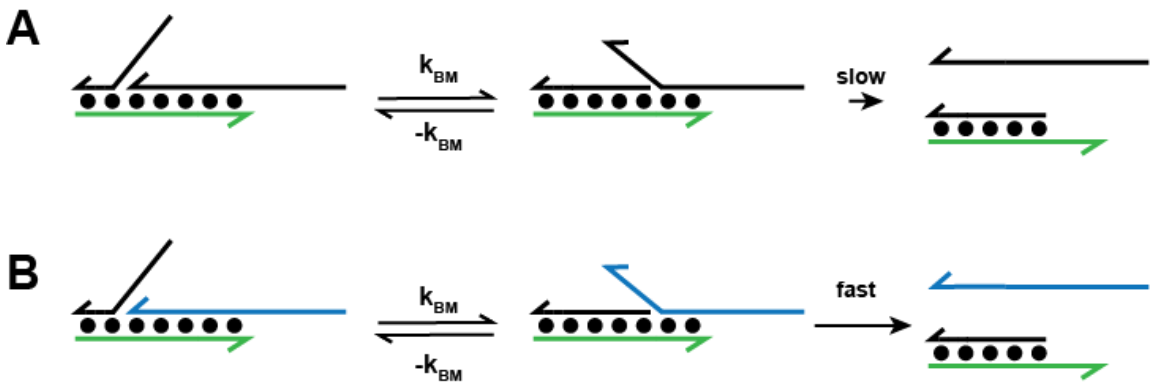



**Figure IV-102. Possible mechanism for the accelerated kinetics observed with truncated inputs in heterochiral strand displacement.** (A) Branch migration of a truncated homochiral input. After branch migration, a 3-way complex is reached that may be long-lived due to the reversibility of branch migration. (B) Branch migration of a truncated heterochiral input. After branch migration, a 3-way complex is reached that may be fast to dissociate rather than branch migrate in the reverse direction. This state mimics the initial 3-way intermediate after a heterochiral input binds to the PNA toehold and may be faster to dissociate due to pathways related to the rate penalty observed in heterochiral strand displacement.

direction. This equilibrium logically explains the reduced rate of strand displacement as you shorten the invader strand, because this concomitantly increases the length of the new toehold which increases the rate of the reverse reaction. However, the process of heterochiral strand displacement suffers from a kinetic penalty regarding strand displacement but can generate product at a reduced rate. It makes sense then, that the reverse reaction *also suffers* from this same penalty, meaning that once branch migration completes the incumbent strand is more likely to dissociate and react with the reporter complex (Figure IV-12 B). However, every strand displacement in this reaction mixture generates a new product duplex with a free toehold (enabling the reverse reaction) that can compete with the reporter for the output strand. This could act as internal negative feedback on the generation of fluorescence and may explain why the heterochiral system with the -8 input does not reach completion during this time course.

### *4.1.4 Studying the kinetic penalty of heterochiral strand displacement*

In traditional DNA strand displacement reactions, it is generally accepted that the overall reaction rate is limited by the rate of branch migration[13]. This is primarily due to a substantial thermodynamic barrier opposing the initiation of branch migration arising from the steric impacts of generating a new single stranded overhang at the branch migration junction as the incumbent strand begins to be displaced. Although toehold nucleation is very fast, it is a reversable process, and this barrier to branch migration increases the likelihood of toehold dissociation rather than successful branch migration. After seeing consistently reduced strand displacement rates when using heterochiral inputs, we were

interested in trying to determine which step during the heterochiral strand displacement pathway was affected. Since fraying of the terminal base pairs of a DNA/DNA duplex (and likely a PNA/DNA duplex) is predicted to occur much faster than the steps of branch migration[13], we focused on the remaining three steps of strand displacement: toehold nucleation and zippering, initiation of branch migration, and branch migration (see steps 1, 3, and 4 from Figure IV-1). We hypothesized that if reaction conditions could be chosen such that toehold nucleation was essentially irreversible, this step could be isolated from initiation and completion of branch migration as well (Figure IV-13 A). If toehold nucleation was irreversible (i.e. the input strand would not dissociate once bound), then it would simply be a matter of *how long* branch migration takes to occur rather than *if* branch migration occurs. Under such experimental conditions, if both the D-A_$L_{kinetics}$ and L-A_$L_{kinetics}$ duplexes were present (along with the corresponding D and L-$R_{kinetics}$) in equal amounts, limiting concentrations of the D-input strand would bind to either the homochiral or heterochiral depending on the kinetics of toehold nucleation. If there is no difference in the kinetics of toehold nucleation between the homochiral and heterochiral reaction pathways, we could expect the input to bind to both stereoisomers of the A duplex to the same extent and each pathway should reach the same endpoint signal at different rates (due exclusively to differences in branch migration between a homochiral or heterochiral branch migration domain, Figure IV-13 B). However, if the rates of toehold nucleation are NOT identical between these two pathways, then a larger fraction of the input would nucleate onto the 'preferred' toehold (most likely the homochiral toehold, based on the

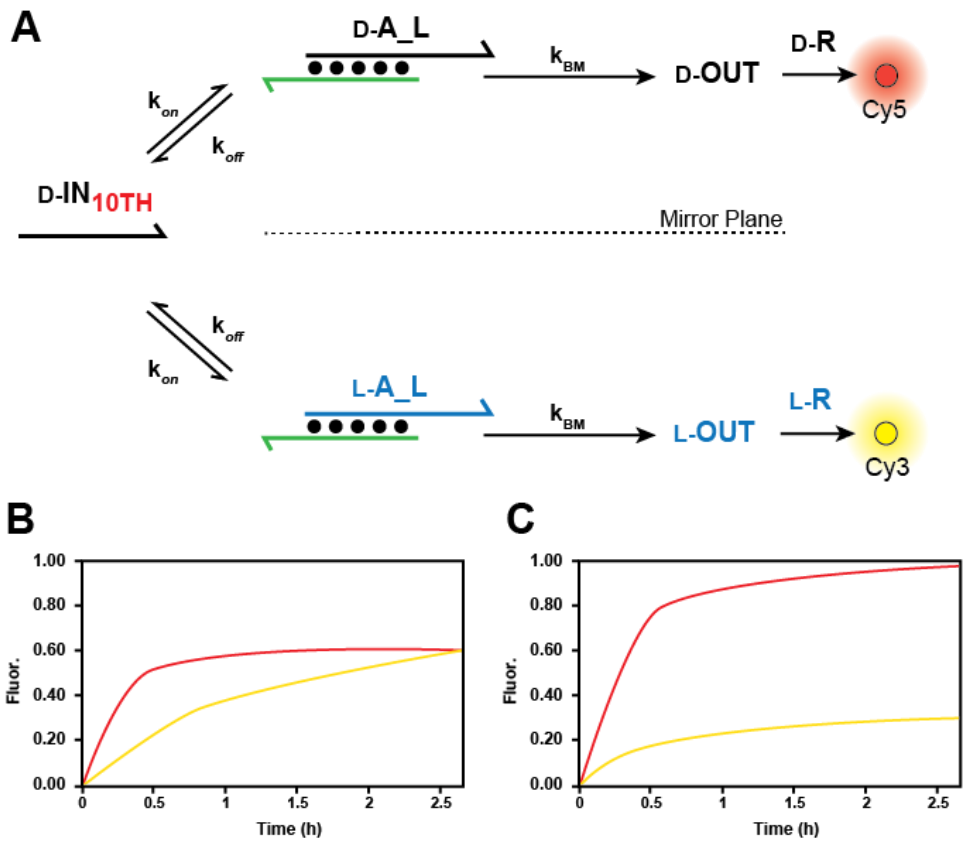



**Figure IV-113. Isolating the kinetic penalty of heterochiral strand displacement.** (A) Kinetic model of strand displacement with both D-A_L$_{kinetics}$ and L-A_L$_{kinetics}$ present in the reaction. (B) If the toehold is long enough that $k_{off}$ is effectively 0, and there is no difference in $k_{on}$ between the heterochiral and homochiral pathways, this represents a potential model for reaction progression. (C) If the toehold is long enough that $k_{off}$ is effectively 0, and there is a difference in $k_{on}$ between the heterochiral and homochiral pathways, this represents a potential model for reaction progression.

current body of data discussed in this thesis), and the preferred pathway would reach a higher signal at the end of the time course (Figure IV-13 C).

The use of PNA in these strand displacement reactions provides me with an opportunity to test this, since PNA/DNA hybridization is more thermodynamically stable than traditional DNA/DNA hybridization. The 10TH input described in section 4.1.1 has a predicted Tm of ~49 ºC, while this might not be completely accurate due to the conditions under which PNA/DNA $T_m$s are predicted[19], it is more than 10 ºC higher that the temperature at which these reactions are run. This suggests that binding of the 10 TH input to the PNA should be essentially irreversible. Strand displacement reactions were set up similar to those described in section 2.1.2, where a thresholding reaction based on the rate difference between the homochiral and heterochiral reactions was discussed. A new version of D-$R_{kinetics}$ was prepared with Cy5 instead of Cy3 so its activation could be monitored simultaneously with L-$R_{kinetics}$, and a single reaction mixture containing 120 nM of the 10 TH input and 300 nM of each reporter was prepared. The reaction was initiated by the addition of 100 nM of both the D-A_$L_{kinetics}$ and L-A_$L_{kinetics}$ duplexes (final concentration) and monitored at 37 ºC for 3 hours. Within this concentration regime the input cannot maximally activate both reporters, and the ratio of Cy5/Cy3 signal will provide insight into the distribution of toehold binding events. Under these conditions, the maximum Cy5/Cy3 ratio should be ~5, assuming all 100 nM of D-$A_{kinetics}$ gets activated before the remaining input activates 20 nM of L-$A_{kinetics}$. This result would indicate that either: toehold binding was still reversible and an increased penalty for heterochiral branch migration funnels input to the homochiral pathway, or that nucleation and zippering are

indeed faster in the homochiral context. The minimum Cy5/Cy3 ratio should be 1, assuming toehold binding is irreversible, and the achiral PNA toehold allows for the 10 TH input to bind equivalently to either D-A_$L_{kinetics}$ or L-A_$L_{kinetics}$. In this case, both reactions should reach the same endpoint signal but at different rates due to the impact of branch migration.

This reaction was first performed at 37 ºC, yielding an observed Cy5/Cy3 ratio of ~2.5 (Figure IV-14 A), potentially indicating a strong preference for binding the toehold of the D-A_$L_{kinetics}$ over the toehold of L-A_$L_{kinetics}$. Such a high ratio was surprising, as it isn't immediately clear why the achiral PNA toehold would display this effect. Due to the potential inaccuracies in PNA/DNA $T_m$ calculations it is possible that toehold binding might still be reversible at 37 ºC. The same experiment was repeated at 30 ºC, almost 20 ºC below the predicted $T_m$ for the 10 TH input to further stabilize toehold binding. This time the Cy5/Cy3 ratio was ~1.6, which still indicates a kinetic preference for the homochiral toehold either through increased association or decreased dissociation compared to the heterochiral toehold (Figure IV-14 B).

As a final test to demonstrate the difference in toehold binding between homochiral and heterochiral systems, a new reporter was designed to isolate toehold binding and branch migration (Figure IV-15 A). This reporter is comprised of D-DNA and has the same 10 nt toehold as the 10 TH input, but the branch migration domain has been removed and replaced with a short stretch of 5 dT bases. This $(dT)_5$ domain was used to mimic the effects of the flexible single-stranded domain of the input strand at the branch migration junction while being unable to participate in branch migration. The 3′ end of the probe

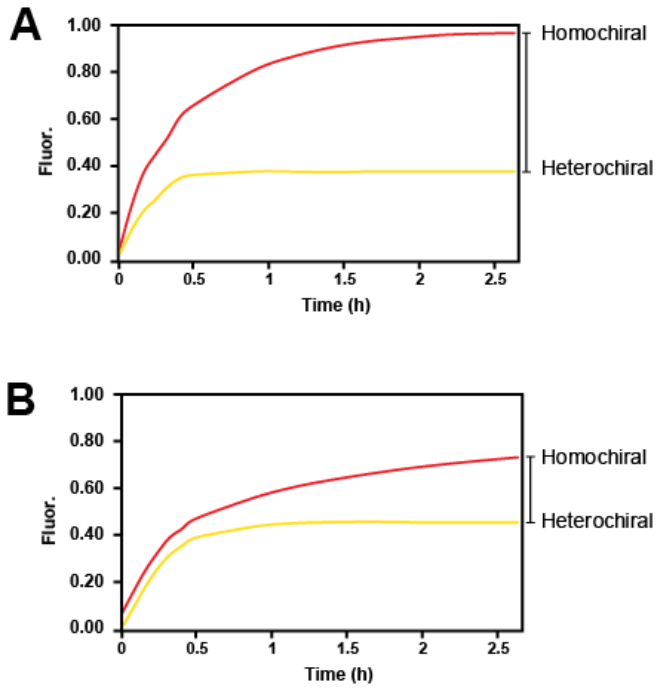



**Figure IV-124. The kinetic penalty of the heterochiral reaction.** In each case, the 10 TH input was present in limiting concentrations. The difference in achieved signal in both experiments suggests that the input 'prefers' its homochiral toehold rather than the heterochiral one. (A) Competition experiment performed at 37 ºC. (B) Competition experiment performed at 30 ºC.

was modified with Black Hole Quencher 2, and the 5′ end of the probe with Cy3. Probes of this design form non-fluorescent complexes due to hydrophobic interactions between the fluorophore and the quencher, but hybridization to its complementary strand disrupts this interaction and generates fluorescence[14, 20]. While my previous experiments relied on the assumption that toehold binding was irreversible under the experimental conditions, this style of reporter should reflect the equilibrium of toehold binding for the homochiral and heterochiral pathways. Briefly, I incubated 100 nM of the toehold probe with 200 nM of either the D-A_$L_{kinetics}$ or L-A_$L_{kinetics}$ duplex at 30 ºC and measured the fluorescence over ~3 hours. As expected, the fluorescence for both the homochiral and heterochiral systems rapidly reached equilibrium (although there was some drift in the fluorescence, this may be due to temperature or instrument fluctuations), indicating the fraction of probe bound to the toehold vs the fraction of unbound probe (Figure IV-15 B). The equilibrium fluorescence of the homochiral pathway was higher than that of the heterochiral pathway, further indicating a higher affinity for the homochiral toehold. Indeed, the ratio of homochiral/heterochiral signal was ~1.35, similar to the 1.6 Cy5/Cy3 ratio identified previously.

This consistent preference for the homochiral reaction pathway is very interesting and might imply that the PNA toehold is partially preorganized due to the chirality of the branch migration domain. Preorganization of the toehold would potentially have a kinetic impact on nucleation and zippering of an input strand (i.e. an increased $k_{on}$), this could explain the apparent preference for the homochiral reaction. This possibility is partially supported in the literature, where it has been shown that PNA strands containing 1-3 chiral

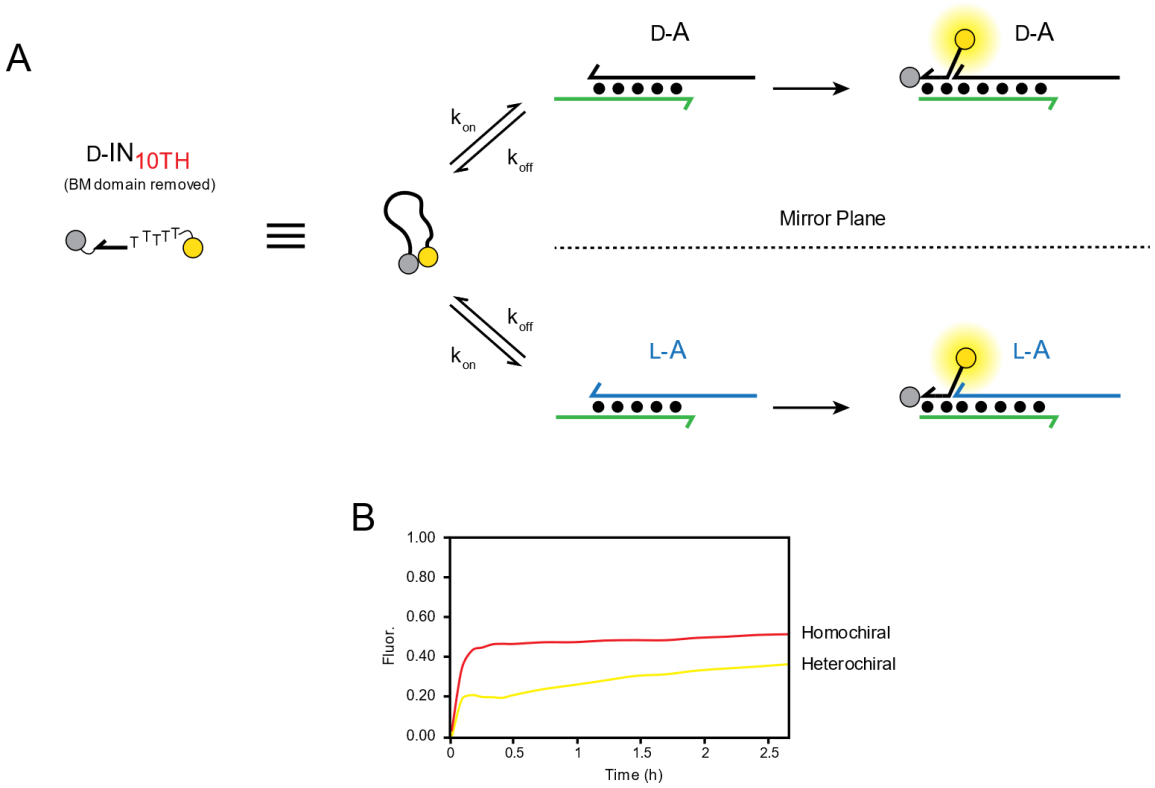



**Figure IV-135. Kinetics of binding a homochiral vs. heterochiral toehold.** (A) A beacon-style reporter complementary to the 10 nt toehold on PNA_$L_{kinetics}$. This reporter cannot undergo branch migration, and its signal should represent the equilibrium of toehold binding. (B) Fluorescence traces demonstrating reporter binding to the homochiral toehold (red) or the heterochiral toehold (yellow).

subunits have faster hybridization kinetics to DNA strands of the same chirality[21]. It has been demonstrated that chirality can been seeded in supramolecular structures by the additional of chiral domains via the 'sergeants and soldiers' and 'majority rules' effects[22] indicating that the chiral branch migration domain could potentially pre-order the PNA toehold as hypothesized here.

### 4.2 Conclusions

While the kinetics and thermodynamics of PNA/DNA hybridization have been studied in some detail, there has been no systematic approach to adapt our knowledge to strand displacement-based systems especially in the context of heterochiral strand displacement. Indeed, there have been very few reports regarding the displacement of a DNA strand from a PNA via toehold mediated strand displacement in the literature[23]. The work that I've described in this chapter improves our understanding of homo- and heterochiral strand displacement reactions from a PNA substrate and will inform the design of novel probes for molecular computation systems and biosensor development. Additionally, I've shown that strand displacement from a PNA/DNA duplex recreates the toehold-dependent exponential acceleration of strand displacement (Figures IV-3 and IV-4), albeit to a lesser extent. I screened a series of input strands that resulted in mismatches within the PNA/DNA waste duplex. This data indicated that heterochiral strand displacement strand displacement strongly discriminates against mismatches within the toehold and near the branch migration domain. This may be an invaluable property for the design of strand displacement probes capable of discriminating against SNPs in cellular

RNAs. Truncated invader strands displayed unique reaction kinetics compared to all-DNA systems. Finally, at the salt and temperature conditions used in these experiments, the 10 TH inputs are expected to be strong enough to effectively not dissociate. Once hybridized to the PNA toehold in these conditions, branch migration should be inevitable, implying that the difference between homochiral and heterochiral reporter activation seen in section 4.1.4 is likely due to a chirality-dependent preference in toehold nucleation. Future experiments are planned to verify these observations. Taken together, this work outlines the potential for heterochiral strand displacement systems much faster than we have previously reported and suggests that such systems may inherently have a high capability for mismatch discrimination. Coupled with the inherent stability of L-DNA based systems, I believe that the chiral inversion step may provide previously unanticipated advantages in the design of nucleic acid probes operating in live cells.

## 4.3 Materials and Methods

### *4.3.1 DNA Design, synthesis and purification*

DNAs and RNAs were either purchased from Integrated DNA Technologies or prepared by solid-phase synthesis on an Expedite 8909 DNA/RNA synthesizer. DNA synthesis reagents, fluorophore phosphoramidites and D-nucleoside phosphoramidites were purchased from Glen Research, and L-nucleoside phosphoramidites were purchased from ChemGenes. Black Hole Quencher 2 (BHQ2) and Black Hole Quencher 3 (BHQ3) CPG resins were purchased from LGC Biosearch technologies. PNAs were purchased from either PNA Bio Inc. or Panagene.

### *4.3.2 Sequence Design*

DNA sequences for the strand displacement circuits depicted in this chapter were based on previous work described by Zhang and Winfree[10]. Mismatch positions were chosen to cover a broad range of positions within the toehold and the branch migration domain.

### *4.3.3 Oligonucleotide purification*

After purchase from IDT, or after all modifications were finished in house, all oligonucleotides were purified by 20% denaturing polyacrylamide gel electrophoresis (PAGE; 19:1 acrylamide:bisacrylamide). Purified ONs were excised from the gel with a sterile razor blade and eluted overnight at room temperature (RT) in a buffer containing 200 mM NaCl, 10 mM EDTA, and 10 mM Tris (pH 7.6). The solution was filtered to remove gel fragments and eluted oligonucleotides were precipitated with ethanol, resuspended in a small volume of ddH2O, and quantified by their absorbance at A260. All 3′ labeled oligonucleotides were synthesized using commercial CPG resins functionalized with the desired modification (e.g. BHQ2) and purified as described above. Fluorescent dyes were coupled directly to the 5′ end as their phosphoramidite using a 5-minute coupling protocol.

### *4.3.4 Preparation of duplex components*

PNA/D-DNA and PNA/L-DNA duplexes were mixed in a 1:1.5 ratio and annealed at 90 ºC in a buffer containing 300 mM NaCl, 1 mM EDTA, 10 mM Tris (pH 7.6). Both stereoisomers of reporter duplexes were annealed in the same fashion, and all duplexes were purified by 20% native PAGE (19:1 acrylamide:bisacrylamide). After purification, each duplex was excised with a sterile scalpel, crushed into a 15 mL conical tube, and eluted in 3 mL of annealing buffer.

### *4.3.5 Monitoring of strand-displacement reactions by spectrofluorimetry*

Each Strand displacement reaction was monitored using a GloMax Discover multi-well plate reader (Promga Crop). All reactions contained 150 nM of the desired input, 100 nM D or L-$A_{kinetics}$ (either the long or short branch migration domain), 300 nM L-R, 300 mM NaCl, 1 mM EDTA, and 10 mM Tris (pH 7.6). Reactions were carried out in a 384-well microplate at a final volume of 30 µL and a temperature of 37 °C unless otherwise specified. Reactions were initiated by the addition of 10 µL of D or L-$A_{kinetics}$ duplex at the appropriate concentration. The fluorescence intensity of reporters labeled with Cy3 was monitored with excitation/emission wavelengths at 520 nm/580-640 nm (bandpass filter), while the fluorescence intensity of the reporters labeled with Cy5 was monitored with excitation/emission wavelengths at 627 nm/660-720 nm (bandpass filter).

All strand displacement reactions were normalized to the signal from a pre-opened reporter representing the maximum achievable fluorescence using the following equation:

$$F_n = \frac{F - F_0}{F_c - F_0}$$

Where Fn is the normalized fluorescence intensity, F is the measured fluorescence at each time point, $F_0$ is the quenched fluorescence, and $F_c$ is the control fluorescence at each time a measurement was taken. This normalization equation allows for us to account for the loss in signal due to photobleaching and enables direct comparison of the different fluorophores used in this study.

### *4.3.6 Monitoring of strand-displacement kinetics by spectrofluorimetry*

For all reactions, I observed a substantial difference in the rate of strand-displacement based on the chirality of the input strand. I determined the rate constant for each of these strand displacement reactions as previously reported[24]. Briefly, the strand displacement reactions were prepared and initiated as described in the previous section and monitored by spectrofluorimetry. Under these conditions, kinetics of reporter opening does not limit the overall reaction rate. All kinetics reactions were performed with 1.5 equivalents of the desired input strand relative to the inversion gate in order to ensure complete displacement of the output strand. The fluorescence curves obtained were fit using an equation derived from the second order rate law with respect to the input strand and the inversion gate. Due the stability of the waste duplexes, all reverse reactions were considered negligible.

## 4.4 References

1. Seeman, N. C.; Sleiman, H. F., DNA nanotechnology. *Nature Reviews Materials* **2017,** *3*, 17068.

2. Seelig, G., et al., Enzyme-Free Nucleic Acid Logic Circuits. *Science* **2006,** *314* (5805), 1585.

3. Zhang, D. Y., et al., Engineering Entropy-Driven Reactions and Networks Catalyzed by DNA. *Science* **2007,** *318* (5853), 1121.

4. Qian, L.; Winfree, E., Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **2011,** *332* (6034), 1196.

5. Srinivas, N., et al., Enzyme-free nucleic acid dynamical systems. *Science* **2017,** *358* (6369), eaal2052.

6. Wan, W. B., et al., Synthesis, biophysical properties and biological activity of second generation antisense oligonucleotides containing chiral phosphorothioate linkages. *Nucleic Acids Res.* **2014,** *42* (22), 13456-13468.

7. McTigue, P. M., et al., Sequence-Dependent Thermodynamic Parameters for Locked Nucleic Acid (LNA)−DNA Duplex Formation. *Biochemistry* **2004,** *43* (18), 5388-5405.

8. Majlessi, M., et al., Advantages of 2′-O-methyl oligoribonucleotide probes for detecting RNA targets. *Nucleic Acids Res.* **1998,** *26* (9), 2224-2229.

9. Kabza, A. M., et al., Heterochiral DNA Strand-Displacement Circuits. *J. Am. Chem. Soc.* **2017,** *139* (49), 17715-17718.

10. Zhang, D. Y.; Winfree, E., Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.* **2009,** *131* (47), 17303-17314.

11. Yurke, B.; Mills, A. P., Using DNA to Power Nanostructures. *Genetic Programming and Evolvable Machines* **2003,** *4* (2), 111-122.

12. Moszyńska, A., et al., SNPs in microRNA target sites and their potential role in human disease. *Open Biol.* **2017,** *7* (4), 170019.

13. Srinivas, N., et al., On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.* **2013,** *41* (22), 10641-10658.

14. Machinek, R. R. F., et al., Programmable energy landscapes for kinetic control of DNA strand displacement. *Nat. Commun.* **2014,** *5* (1), 5324.

15. Broadwater, D. W. B.; Kim, Harold D., The Effect of Basepair Mismatch on DNA Strand Displacement. *Biophys. J.* **2016,** *110* (7), 1476-1484.

16. Eriksson, M.; Nielsen, P. E., PNA-nucleic acid complexes. Structure, stability and dynamics. *Q. Rev. Biophys.* **1996,** *29* (4), 369-394.

17. Igloi, G. L., Variability in the stability of DNA–peptide nucleic acid (PNA) single-base mismatched duplexes: Real-time hybridization during affinity electrophoresis in PNA-containing gels. *Proc. Natl. Acad. Sci.* **1998,** *95* (15), 8562.

18. Yuan, B.-f., et al., Kinetics of base stacking-aided DNAhybridization. *Chem. Commun.* **2008,** (48), 6600-6602.

19. Giesen, U., et al., A formula for thermal stability (Tm) prediction of PNA/DNA duplexes. *Nucleic Acids Res.* **1998,** *26* (21), 5004-5006.

20. Yurke, B., et al., A DNA-fuelled molecular machine made of DNA. *Nature* **2000,** *406* (6796), 605-608.

21. Menchise, V., et al., Insights into peptide nucleic acid (PNA) structural features: The crystal structure of a D-lysine-based chiral PNA–DNA duplex. *Proc. Natl. Acad. Sci.* **2003,** *100* (21), 12021.

22. Morrow, S. M., et al., Transmission of chirality through space and across length scales. *Nat. Nanotechnol.* **2017,** *12*, 410.

23. Wang, Z., et al., Imaging mRNA Expression in Live Cells via PNA/DNA Strand Displacement-Activated Probes. *J. Nucleic Acids* **2012,** *2012*, 11.

24. Olson, X., et al., Kinetics of DNA Strand Displacement Systems with Locked Nucleic Acids. *J. Phys. Chem. B* **2017,** *121* (12), 2594-2602.

# CHAPTER V

# SUMMARY AND OUTLOOK

## 5.1 Summary

The work presented in this dissertation outlines the discovery and partial characterization of novel DNA strand displacement reactions enabling, for the first time, the sequence specific conversion of a D-DNA input into an L-DNA output. Chapter 1 provides a general overview of DNA and RNA, describing their structure, biological function and disease diagnosis and treatment. Common methods for the sequence specific detection of nucleic acids *in vitro* and *in vivo*, and the advent of dynamic DNA-based devices having the potential to build much more complex systems, are discussed. Importantly, this chapter establishes the need for new biostable DNA devices capable of functioning in the complex cellular milieu.

### *5.1.1 Development of heterochiral strand displacement reactions*

Chapter 2 focused on the invention of two novel strand displacement reactions: one employing a PNA/DNA heteroduplex with an achiral toehold, and the other showing that chimeric D-DNA and L-DNA complexes could be fine-tuned such that disruption of the D-DNA region would cause the L-DNA region to spontaneously melt. Both reactions enable a D-DNA invader based on the miR-155 sequence to directly release an L-DNA output. The kinetics of each displacement reaction are discussed, and we observed an interesting penalty with regards to displacing an incumbent strand of one chirality from a PNA/DNA duplex via invasion of a heterochiral input. Both systems were able to detect

the RNA versions of their targets even in the presence of total RNA lysate, establishing their potential for the development of a novel class of nuclease-resistant biosensors. To prove the generality of design 1 we prepared 2 additional systems capable of specifically detecting miR-10b as well as the MnSOD mRNA. In an effort to increase the complexity of these systems, the first L-DNA AND gate was designed based on the inversion gates for miR-155 and miR-10b. In this design, both target sequences need to be present in order to release a fluorophore from a quencher demonstrating Boolean AND logic.

*5.1.2 Validation of heterochiral strand displacement reaction performance in live cells*

In Chapter 3 we ordered a fully 2′-OMe RNA version of the 155 system and tested its function *in vitro* as a point of comparison between a standard modified DNA and our L-DNA devices. Interestingly, this system behaved poorly *in vitro*, showing much slower activation than either the homo- or heterochiral displacement from PNA. This highlights the comparative simplicity of our design compared to other DNA/RNA modifications. All 3 miR-155 reporter variants (D-DNA, L-DNA and 2′-OMe RNA) were transfected into live cells with or without their inputs to test their stability and activation in the cellular environment. Input specific activation of the L-reporter increased up to 6 hours, while the reporter alone did show a minor increase in background signal. The D-reporter, however, achieved a lower maximum fluorescence in cells with a much higher background. This is likely due to a combination of input degradation (preventing activation) and reporter degradation (leading to high background), a common probably facing unmodified DNA devices in live cells. The 2′-OMe reporter was surprisingly stable in the cellular

environment, with almost no degradation observed even over 24 hours. This could be due in part to the increased $T_m$ of fully modified 2′-OMe duplexes, decreasing the likelihood of duplex melting due to non-specific interactions in the cell. Unfortunately, this reporter was poorly activated as well, mirroring its performance *in vitro*. Additionally, while 2′-OMe RNAs are very resistant to endonuclease degradation within duplex regions, they are susceptible to exonuclease degradation which could degrade both the single stranded input and the single stranded toehold on the reporter. This would also manifest as low activation inside cells. Then, two single component heterochiral strand displacement systems (155 and 10b) were transfected into cells with either the correct D or L-input, or a version of the input with a scrambled toehold. This established that each single component heterochiral strand displacement was activated in a toehold-dependent manner. The two component L-AND gate was transfected similarly, and in all cases the L-systems showed higher stability and percent activation that their D-counterparts. This work establishes heterochiral strand displacement as a potential route for the design of biostable probes for the detection of one or more cellular nucleic acid inputs.

### *5.1.3 Characterizing the kinetics of heterochiral strand displacement reactions employing a PNA/DNA heteroduplex*

Chapter 4 focused on the development of a better understanding of heterochiral strand displacement systems utilizing a PNA/DNA heteroduplex. Importantly, we sought some understanding regarding the dramatic rate decrease observed between homochiral and heterochiral inputs. For this, we designed a new system based on work by Zhang and

Winfree[1] to systematically test a series of input strands and evaluate their impact on the rate of strand displacement. Using a set of inputs progressively truncated within the toehold region, the impact of toehold length on the rate of strand displacement was characterized for both the homo- and heterochiral inputs. Then, we tested a series of input strands containing mismatches throughout the toehold and branch migration domains. Since PNA is known to strongly discriminate against mismatches, we expected that the kinetics of heterochiral strand displacement would be strongly affected by mismatches within the input strand. This does seem to be the case, as the only tolerated mismatches were those that appeared late in the branch migration domain. While this trend is consistent with what is seen in all-DNA strand displacement reactions, the inherent rate decrease in heterochiral strand displacement renders these mismatches virtually incapable of strand displacement. This may prove to be a powerful benefit in the use of this style heterochiral strand displacement reaction. Finally, we tested a series of inputs progressively truncated at the end of the branch migration domain to evaluate the utility of this style heterochiral strand displacement for 'toehold exchange' type reactions[1-2]. This type of strand displacement reaction reveals a new toehold after the incumbent strand is displaced, which is often used to enable alternative functions with the strand displacement regime. Alternatively, if the incumbent strand can be efficiently displaced without the need for the input to migrate fully through the branch migration region, reporter modules could be designed independent of target sequence, increasing the modularity of these new strand displacement systems. Inputs truncated by 4 (-4) or 6 (-6) bases showed very little difference in reaction rate when compared to a full-length input having the same length

toehold. Interestingly, inputs truncated by 8 (-8) bases substantially slowed the reaction of the homochiral system. However, the -8 input rapidly activated the heterochiral system to about 50% of the max signal before plateauing. Even though it didn't reach max signal, this is the only example in this body of experiments where the heterochiral system was faster than the homochiral one. This phenomenon was reproducible in triplicate samples, so it will be interesting to pursue this observation further.

### 5.2 Outlook

The work outlined in section 6.1 demonstrates, for the first time, the sequence specific transfer of information across the chiral mirror. Since the unnatural enantiomer of D-DNA, L-DNA, is highly resistant to nuclease degradation and off target interactions this work begins to address some pressing concerns in biosensor design and smart therapeutics. Indeed, work in our lab is ongoing with this style of strand displacement and we've recently shown the sequence specific detection of a microRNA target artificially upregulated in HeLa cells using a single component inversion gate releasing an RNA strand based on the mango aptamer[3]. Additionally, our lab is attempting to use L-DNA inputs to exert sequence-dependent control over D-DNA based therapeutics inside live cells. The work described in this thesis and the ongoing work in our lab has identified a few hurdles to address: 1) Scaling up complexity of these systems, 2) increasing signal to noise in live cells in particularly by decreasing input-independent leak, and 3) delivery of nuclease resistant circuit components into live cells.

### *5.2.1 Scaling up the complexity of heterochiral strand displacement systems in live cells*

While we've described the general design of a PNA/DNA heterochiral strand displacement system, and conveniently applied that design to develop a biostable L-AND gate, there is more to be done in the development and testing of L-DNA molecular sensors. The work in chapter 4 begins to expand our understanding of these heterochiral strand displacement reactions regarding toehold acceleration, mismatch discrimination, and input length dependence. These properties will play an important role in designing new modules for the detection of cellular RNAs, increasing the rate of strand displacement and reduce spurious leak reactions. One of the most valuable functions enabled by DNA computation is signal amplification. These techniques (i.e. fuel strand catalysis, hybridization chain reaction, or catalytic hairpin assembly) allow for the conversion of a small amount of one input signal to a large amount of a fluorescent signal. Many of the previous techniques can generate multi-fold amplification *in vitro*[4-6] and have been applied with limited success in live cells[7-9], but have yet to be realized in a heterochiral fashion. Our lab has developed a few early designs *in vitro* but attempts to package them in transfection reagents and activate them in live cells have faced issues with spurious opening and high background mitigating the benefit of such systems.

### *5.2.2 Increasing signal to noise in live cells*

Integral to the design of any biosensor or smart therapeutic triggered by a specific set of inputs, is the ability to remain OFF in the absence of those inputs. Conversely, once the inputs have been registered and the system is converted to an ON state, the change in

signal must be readily distinguishable from the background. This is primarily achieved through the use of fluorophore – quencher pairs in DNA reporter designs, and the same is true of the work described herein. This is an issue that has been essentially solved with DNA-based catalytic systems in vitro[1-2], but achieving the same level of amplification without a concomitant increase in leak due to degradation or environmental factors has been difficult in living systems. While the initial application of this research focused on increasing signal to noise through the use of biostable L-DNA reporters there are other design parameters to consider: fluorophore – quencher pair selection may provide better quenching or higher fluorescence in the cellular environment. Design of ratiometric systems would allow for the normalization of signal intensity to probe delivery, which should help reduce experimental variability.

*5.2.3 Delivery of L-DNA circuit components into live cells*

As with nearly every DNA based probe or therapeutic, the delivery of these heterochiral strand displacement devices has proven to be one of the major challenges in our hands. While we focused on the use of L2000 in the experiments described in this work, our work with transfection reagents has been less than consistent. Very few of the transfection reagents in the panel show in Figure III-1 were even able to deliver these small nucleic acids and L2000, which was the best by a large margin, experienced relatively large batch to batch variation with regards to delivery and preventing pre-interaction of circuit components. Additionally, some lots of L2000 seemed to strongly deliver these components into the nucleus, dramatically decreasing its utility as an agent

for the delivery of biosensors targeting nucleic acids localized elsewhere in the cell. Our lab has explored a variety of alternative delivery methods, with some success such as electroporation, small molecule delivery enhancement (cholesterol, folate, GalNAC, etc.), and aptamer targeted delivery.

### 5.3 References


1. Zhang, D. Y.; Winfree, E., Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.* **2009,** *131* (47), 17303-17314.

2. Qian, L.; Winfree, E., Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **2011,** *332* (6034), 1196.

3. Zhong, W.; Sczepanski, J. T., A Mirror Image Fluorogenic Aptamer Sensor for Live-Cell Imaging of MicroRNAs. *ACS sens.* **2019,** *4* (3), 566-570.

4. Yin, P., et al., Programming biomolecular self-assembly pathways. *Nature* **2008,** *451* (7176), 318-322.

5. Zhang, D. Y., et al., Engineering Entropy-Driven Reactions and Networks Catalyzed by DNA. *Science* **2007,** *318* (5853), 1121.

6. Dirks, R. M.; Pierce, N. A., Triggered amplification by hybridization chain reaction. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101* (43), 15275.

7. He, L., et al., mRNA-Initiated, Three-Dimensional DNA Amplifier Able to Function inside Living Cells. *J. Am. Chem. Soc.* **2018,** *140* (1), 258-263.

8. Wu, C., et al., A Nonenzymatic Hairpin DNA Cascade Reaction Provides High Signal Gain of mRNA Imaging inside Live Cells. *J. Am. Chem. Soc.* **2015,** *137* (15), 4900-4903.

9. Cheglakov, Z., et al., Live Cell MicroRNA Imaging Using Cascade Hybridization Reaction. *J. Am. Chem. Soc.* **2015,** *137* (19), 6116-6119.

# APPENDIX A

## DNA, RNA AND PNA SEQUENCES

| Name | Sequence Identity 5′→3′ | Makes complex |
|---|---|---|
| PNA$_{155}$ | CCTATCACGATTAGCATTAA | A$_{155}$ |
| L-OUT$_{155}$ | CTAATCGTGATAGGATCGAACTGGTACG | L-A$_{155}$ |
| L-OUT$_{155}$ | CTAATCGTGATAGGATCGAACTGGTACG | D-A$_{155}$ |
| D-IN$_{155}$ | TTAATGCTAATCGTGATAGG | n/a |
| L-IN$_{155}$ | TTAATGCTAATCGTGATAGG | n/a |
| D-RNA$_{155}$ | UUAAUGCUAAUCGUGAUAGGGU | n/a |
| D-F$_{155}$ | /5Cy5/GGCGTACCAGTTCGATCCTATC | D-R$_{155}$ |
| D-Q$_{155}$ | ATCGAACTGGTACGCC/3BHQ3/ | |
| L-F$_{155}$ | /5Cy5/GGCGTACCAGTTCGATCCTATC | L-R$_{155}$ |
| L-Q$_{155}$ | ATCGAACTGGTACGCC/3BHQ3/ | |
| D-F$_{155.2}$ | /5Cy3/GGCGTACCAGTTCGATCCTATCAC | D-R$_{155.2}$ |
| D-Q$_{155.2}$ | GGATCGAACTGGTACGCC/3BHQ3/ | |
| L-F$_{155.2}$ | /5Cy5/GGCGTACCAGTTCGATCCTATCAC | L-R$_{155.2}$ |
| L-Q$_{155.2}$ | GGATCGAACTGGTACGCC/3BHQ3/ | |
| C1 | AATCGTGATAGGGGTGTGATAGGATCGAACTGGTACG | D/L-A$_1$ |
| C2 | ATCCTATCACACCCCTATCACGATTAGCATTA | |
| C3 | AATCGTGATAGGGGTATCCTATCAC | D/L-A$_2$ (with L-OUT$_{155}$) |
| C4 | ACCAGTTCGACCCCTATCACGATTAGCATTAA | |
| PNA$_{10b}$ | CACAAATTCGGTTCTACAG | A$_{10b}$ |
| L-OUT$_{10b}$ | AACCGAATTTGTGCAAAGCCCATCCCACC | L-A$_{10b}$ |
| L-OUT$_{10b}$ | AACCGAATTTGTGCAAAGCCCATCCCACC | D-A$_{10b}$ |
| D-IN$_{10b}$ | TACCCTGTAGAACCGAATTTGTG | n/a |

**Table A - 1.** D-DNA, L-DNA and PNA sequences discussed in Chapter 2.

| Name | Sequence Identity 5′→3′ | Makes complex |
|---|---|---|
| L-$IN_{10b}$ | TACCCTGTAGAACCGAATTTGTG | n/a |
| D-$RNA_{10b}$ | UACCCUGUAGAACCGAAUUUGUG | n/a |
| D-$F_{10b}$ | /5Cy5/GGTGGGTAGGGCTTTGATTCGG | D-$R_{10b}$ |
| D-$Q_{10b}$ | CAAAGCCCTACCCACC/3BHQ3/ | |
| L-$F_{10b}$ | /5Cy5/GGTGGGTAGGGCTTTGATTCGG | L-$R_{10b}$ |
| L-$Q_{10b}$ | CAAAGCCCTACCCACC/3BHQ3/ | |
| $PNA_{MnSOD}$ | AGTTACATTCTCCCAGTTGATT | $A_{MnSOD}$ |
| L-$OUT_{MnSOD}$ | TAAGAGGGTCAACTAATCAATGCTAACTAA | L-$A_{MnSOD}$ |
| L-$OUT_{MnSOD}$ | TAAGAGGGTCAACTAATCAATGCTAACTAA | D-$A_{MnSOD}$ |
| D-$IN_{MnSOD}$ | AATCAACTGGGAGAATGTAACT | n/a |
| L-$IN_{MnSOD}$ | AATCAACTGGGAGAATGTAACT | n/a |
| D-$F_{MnSOD}$ | /5Cy3/GGTCTAATCAATCGTAACTAATC | D-$R_{MnSOD}$ |
| D-$Q_{MnSOD}$ | CCAGTTGATTAGTTACGATTGATTAGACC/3BHQ2/ | |
| L-$F_{MnSOD}$ | /5Cy3/GGTCTAATCAATCGTAACTAATC | L-$R_{MnSOD}$ |
| L-$Q_{MnSOD}$ | CCAGTTGATTAGTTACGATTGATTAGACC/3BHQ2/ | |
| D-$Block_{AND}$ | TGTGCAAAGCCCATCCCACCGTGATA | D-$R_{AND}$ |
| D-$Q_{AND}$ | GGATCGAACTGGTACGCC/3BHQ2/ | |
| D-$F_{AND}$ | /5Cy3/GGCGTACCAGTTCGATCCTATCACGGTGGGATGGGCTTTGCACAAATTCG | |
| L-$Block_{AND}$ | TGTGCAAAGCCCATCCCACCGTGATA | L-$R_{AND}$ |
| L-$Q_{AND}$ | GGATCGAACTGGTACGCC/3BHQ2/ | |
| L-$F_{AND}$ | /5Cy3/GGCGTACCAGTTCGATCCTATCACGGTGGGATGGGCTTTGCACAAATTCG | |

**Table A - 1.** Continued.

| Name | Sequence Identity 5′→3′ | Makes complex |
|---|---|---|
| $PNA_{155}$ | CCTATCACGATTAGCATTAA | $A_{155}$ |
| L-$OUT_{155}$ | CTAATCGTGATAGGATCGAACTGGTACG | L-$A_{155}$ |
| D-$OUT_{155}$ | CTAATCGTGATAGGATCGAACTGGTACG | D-$A_{155}$ |
| $PNA_{155}$ | CCTATCACGATTAGCATTAA | 2′-OMe-$A_{155}$ |
| 2′-OMe-$OUT_{155}$ | CTAATCGTGATAGGATCGAACTGGTACG | |
| D-$F_{155.2}$ | /5Cy3/GGCGTACCAGTTCGATCCTATCAC | D-$R_{155.2}$ |
| D-$Q_{155.2}$ | GGATCGAACTGGTACGCC/3BHQ3/ | |
| L-$F_{155.2}$ | /5Cy5/GGCGTACCAGTTCGATCCTATCAC | L-$R_{155.2}$ |
| L-$Q_{155.2}$ | GGATCGAACTGGTACGCC/3BHQ3/ | |
| 2′-OMe-$F_{155.2}$ | /5Cy3/GGCGTACCAGTTCGATCCTATCAC | 2′-OMe-$R_{155}$ |
| 2′-OMe-$Q_{155.2}$ | GGATCGAACTGGTACGCC/3BHQ3/ | |
| $PNA_{10b}$ | CACAAATTCGGTTCTACAG | $A_{10b}$ |
| L-$OUT_{10b}$ | AACCGAATTTGTGCAAAGCCCATCCCACC | L-$A_{10b}$ |
| L-$OUT_{10b}$ | AACCGAATTTGTGCAAAGCCCATCCCACC | D-$A_{10b}$ |
| D-$IN_{10b}$ | TACCCTGTAGAACCGAATTTGTG | n/a |
| L-$IN_{10b}$ | TACCCTGTAGAACCGAATTTGTG | n/a |
| D-$F_{10b}$ | /5Cy5/GGTGGGTAGGGCTTTGATTCGG | D-$R_{10b}$ |
| D-$Q_{10b}$ | CAAAGCCCTACCCACC/3BHQ3/ | |
| L-$F_{10b}$ | /5Cy5/GGTGGGTAGGGCTTTGATTCGG | L-$R_{10b}$ |
| L-$Q_{10b}$ | CAAAGCCCTACCCACC/3BHQ3/ | |

**Table A - 2.** D-DNA, L-DNA and PNA sequences discussed in Chapter 3.

| Name | Sequence Identity 5′→3′ | Makes complex |
|---|---|---|
| D-Block$_{AND}$ | TGTGCAAAGCCCATCCCACCGTGATA | D-R$_{AND}$ |
| D-Q$_{AND}$ | GGATCGAACTGGTACGCC/3BHQ2/ | |
| D-F$_{AND}$ | /5Cy3/GGCGTACCAGTTCGATCCTATCACGGTGGGATGGGCTTTGCACAAATTCG | |
| L-Block$_{AND}$ | TGTGCAAAGCCCATCCCACCGTGATA | L-R$_{AND}$ |
| L-Q$_{AND}$ | GGATCGAACTGGTACGCC/3BHQ2/ | |
| L-F$_{AND}$ | /5Cy3/GGCGTACCAGTTCGATCCTATCACGGTGGGATGGGCTTTGCACAAATTCG | |

**Table A - 2.** Continued.

| Name | Sequence Identity 5′→3′ | Makes complex |
|---|---|---|
| PNA_L$_{kinetics}$ | ATGTCACTTCCGTAGGGTATTGAATGAGGG | A_L$_{kinetics}$ |
| PNA_S$_{kinetics}$ | ATGTCACTTCCGTAGGGTATTGAATG | A_S$_{kinetics}$ |
| L-OUT$_{kinetics}$ | CCACATACATCATATTCCCTCATTCAATACCCTACG | L-A_X$_{KINETICS}$ |
| D-OUT$_{kinetics}$ | CCACATACATCATATTCCCTCATTCAATACCCTACG | D-A_X$_{KINETICS}$ |
| D-F$_{kinetics}$ | /5Cy3/CACATCATATTCCCTCATTC | D-R$_{kinetics}$ |
| D-Q$_{kinetics}$ | AGGGTATTGAATGAGGGAATATGATGTG/3BHQ3/ | |
| L-F$_{kinetics}$ | /5Cy5/CACATCATATTCCCTCATTC | L-R$_{kinetics}$ |
| L-Q$_{kinetics}$ | AGGGTATTGAATGAGGGAATATGATGTG/3BHQ3/ | |
| D-10TH$_{beacon}$ | /5Cy3/TTTTTGAAGTGACAT/3BHQ2/ | n/a |
| D-10TH | CCCTCATTCAATACCCTACGGAAGTGACAT | n/a |
| D-8 TH | CCCTCATTCAATACCCTACGGAAGTGAC | n/a |
| D-6 TH | CCCTCATTCAATACCCTACGGAAGTG | n/a |
| D-4 TH | CCCTCATTCAATACCCTACGGAAG | n/a |
| D-2 TH | CCCTCATTCAATACCCTACGGA | n/a |
| D-BM2 | CCGTCATTCAATACCCTACGGAAGTGAC | n/a |
| D-BM1 | CCCTCATTCATTACCCTACGGAAGTGAC | n/a |
| D-1BM | CCCTCATTCAATACCCTACCGAAGTGAC | n/a |
| D-TH2 | CCCTCATTCAATACCCTACGGAACAGAC | n/a |
| D-TH1 | CCCTCATTCAATACCCTACGGAAGAGAC | n/a |
| D-1TH | CCCTCATTCAATACCCTACGCAAGTGAC | n/a |
| D-IN-4 | CATTCAATACCCTACGGAAGTGAC | n/a |
| D-IN-6 | TTCAATACCCTACGGAAGTGAC | n/a |
| D-IN-8 | CAATACCCTACGGAAGTGAC | n/a |

**Table A - 3.** D-DNA, L-DNA and PNA sequences discussed in Chapter 4.

# APPENDIX B

## EXAMPLE NUPACK CODE

```
#
# design material, temperature, and trials
# see NUPACK User Guide for valid options for
# material, sodium, magnesium, and dangles
#
material = dna
temperature[C] = 23.0    # optional units: C (default) or K
trials = 2
sodium[M] = .30    # optional units: M (default), mM, uM, nM, pM
dangles = some

#
# target structure using DU+ notation
#
structure miR = U27

#
# sequence domains
#
domain a = CTGTAGAACCGAA N14
```

```
#
# thread sequence domains onto target structures
#
miR.seq = a


#
# specify stop conditions for normalized ensemble defect
# default: 1.0 (percent) for each target structure
#
miR.stop = 1.0


#
# prevent sequence patterns
#
prevent = AAA, CCC, GGG, UUU, KKKKK, MMMMM, RRRRR, SSSSS,
WWWWW, YYYYY, AAAA, CCCC, GGGG, UUUU, KKKKKK, MMMMMM,
RRRRRR, SSSSSS, WWWWWW, YYYYYY
```