

RICE UNIVERSITY

Multiplexed and Reiterative Detection of Protein Markers in Cells using Dynamic Nucleic Acid Complexes

by

Dzifa Yawa Duose

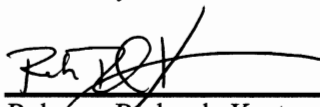
A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE



Michael R. Diehl, Chair
Assistant Professor of Bioengineering and
Chemistry



Rebecca Richards-Kortum
Stanley C. Moore Professor of Bioengineering



Stephan Link
Assistant Professor of Chemistry



Walter N. Hittelman
Professor, Experimental Therapeutics
UTMDACC

HOUSTON, TEXAS

April 2011

Abstract

Multiplexed and Reiterative Detection of Protein Markers in Cells using Dynamic Nucleic Acid Complexes

by

Dzifa Yawa Duose

The diagnosis, staging and clinical management of cancer and other diseases is becoming increasingly reliant upon the identification and quantification of molecular markers as well their spatial distribution in histological samples. Yet, due to spectral overlap of dyes and the inability to remove probes without affecting marker integrity, immunohistological methods are limited by the number of markers that can be examined on a single specimen resulting in loss of information that could be vital to diagnosis or treatment.

This dissertation describes the development and characterization of an erasable multi-color imaging technology capable of detecting large numbers of molecular markers on a single biological sample. The system consists of (1) 'targets', which are single or partially hybridized DNA strands conjugated to a protein of interest for biomarker recognition in cells, and (2) multi-strand, fluorophore-containing DNA 'probe complexes' that react with the DNA portion of the target in a sequence dependent fashion to create fluorescent reporting complexes. The addition of a quencher-bearing ssDNA displaces the target's DNA strand to effectively remove the dye from the marker so that the sample can be re-imaged for other markers with minimal interference from prior

rounds of labeling. Orthogonal DNA sequences and spectrally-separated dyes can be used to create multiple, unique target/probe pairs that associate specifically and can be imaged in parallel.

The overall utility of this technology depends on high specificity of targets to respective probe complexes, highly efficient labeling and erasing to ensure that fluorescent signals can be used to fully quantify target abundance without the interference of signals from previous rounds of labeling, and short reaction times to allow for multiple rounds of processing on the same sample without loss of integrity. Based on the above criteria, three classes of probes were designed and their structure-function relationships elucidated to determine the contributions of complex size, free energy differences between intermediate states, and strand displacement on labeling and erasing kinetics and efficiencies on cells.

A comparison of the kinetics of the labeling and erasing reactions for the three different constructs showed that reaction efficiencies depend less on calculated net free energy change than on the engineered state of the complex during the strand displacement reaction (i.e., the type of strand displacement reaction it participates in). This new paradigm in probe design allowed the system to meet its design goals, potentially increasing the diagnostic power of individual histological specimens and opening the door to more sophisticated analyses of cell phenotype and its functional relationship to disease.

Acknowledgments

I would like to thank the following people without whom this journey would have been difficult to complete; Arthur (Rusty) Rogers, Ryan Schweller and Jan Zimack for being my cohorts in the development and realization of this project. Special thanks to Pamela Constantinou for instilling in me the joy of science and encouraging me to pursue my dreams. I would also like to thank Kenneth Jamison and Jonathan Driver, my confidantes and holiday buddies. I shall always cherish your friendship and support and remember all the fun times we had together.

I want to thank Mike for being a great teacher and advisor and for guiding me from water pipetting exercises to the publication of paper, an industrial internship and the completion of this dissertation. Thank you for making me a good scientist.

I would also like to thank Walter Hittelman, for taking me into his lab and providing me with training and the critical assessment of my work.

Thanks to all my friends and family who have supported me spiritually, emotionally and financially throughout this process. Patricia Kalu, Kofi Ntim, Mawuena Duose and Angela Boateng, I appreciate your being there for me through thick and thin.

Finally, to my Mum and Dad, thanks for having faith in me and for bringing me up to believe that all things are possible with hard work and prayer.

Contents

Acknowledgments	iv
Contents	v
List of Figures	1
1.1. Overview	3
1.2. Chapter Summaries	5
Background	8
2.1. Motivation	8
2.2. The need for multiplexed immuno-based imaging techniques	10
2.2.1. Example Application: Cancer Stem Cell Identification	13
2.3. Current Technologies	15
2.3.1. Development of optics and dyes for multiplexed marker imaging.....	16
2.3.2. Sequential analysis of biological specimen for multiplexed marker imaging. 17	
2.4. The solution: Multiplexed and Reiterative Molecular Biomarker Imaging (MRMBI)	19
2.4.1. Adaptation of Entropy-driven DNA circuits for Multiplexed and Reiterative Molecular Biomarker Imaging	23
Multiplexed and Reiterative Fluorescence Labeling via DNA Circuitry	27
3.1. Introduction	27
3.2. Materials and Methods	28
3.2.1. DNA Circuit Design and Characterization	28
3.2.2. Conjugation of ssDNA catalyst with artificial protein (leucine zipper)	29
3.2.3. Microarray Procedures.....	30
3.2.4. Cell Culture and Labeling.....	31
3.2.5. Fluorescence Imaging and Analyses	32
3.3. Results and Discussion.....	32
3.4. Conclusion.....	45
Optimized On-Cell DNA Strand Exchange Reactions for In Situ Marker Analyses	46
4.1. Introduction	47
4.2. Materials and Methods	49

4.2.1. DNA Probe Design	50
4.2.2. Cell Labeling and Erasing Procedures	52
4.3. Results and Discussion	53
4.3.1. Dynamic DNA Probe Designs	53
4.3.2. Selective <i>In Situ</i> Labeling and Erasing of DNA-conjugated Proteins	57
4.3.3. On-Cell Kinetics of Strand Displacement Reactions	59
4.3.4. Efficient Signal Erasing via Four-way Branched Migration Reactions	65
Conclusion	68
5.1. Summary and Conclusions	68
5.1.1. Summary of DNA Constructs Designed	70
5.2. Future Directions	72
5.2.1. Final probe design	72
5.2.2. Multiplexed imaging with an antibody binding protein: LG	74
5.2.3. Amplification	77
References	80
Appendix A	86

List of Figures

Figure 2.1 : Tissue section of a H460 lung tumor xenograft²¹.	12
Figure 2.2: Identification of tumorigenic breast cancer cells.	14
Figure 2.3: DNA strand displacement mechanism.	21
Figure 2.4: Multiplexed biomarker imaging via Multiplexed and Reiterative Molecular Biomarker Imaging.	23
Figure 2.5: The adaptation of entropically driven circuit for MRMBI.	24
Figure 3.1: DNA circuit-based marker labeling and dye removal reactions. ...	33
Figure 3.2: Native PAGE gel displaying DNA circuit reaction products	34
Figure 3.3: Concentration gradient of catalyst microarray.	36
Figure 3.4. Reiterative Labeling of individual arrayed catalysts.	37
Figure 3.5: Multiplexed and reiterative labeling of multiple DNA catalysts. ...	39
Figure 3.6 Determination of background circuit reactivity	40
Figure 3.7: Targeting DNA circuit complex to GFP-Z_E.	41
Figure 3.8: Circuit-based labeling of GFP-ZE transfected HeLa cells.	42
Figure 3.9: Reiterative circuit labeling of GFP-ZE transfected cells.	44
Figure 4.1: Labeling and Erasing Displacement Reactions using a 2-strand Probe complex.	50
Figure 4.2: Labeling and Erasing Displacement Reactions using a 3-strand Probe	50
Figure 4.3: Labeling and removal of Cy5 fluorophores from protein markers via strand displacement reactions of a 2-strand probe complex.	54
Figure 4.4: Labeling and removal of Cy5 fluorophores from protein markers via strand displacement reactions of a 3-strand probe complex.	55

Figure 4.5: Kinetics of DNA strand Displacement reactions on fixed cells.....	59
Figure 4.6: Intermediate Complex Formation during Erasing Displacement Reactions with a 3-strand probe complex.....	61
Figure 4.7: Labeling and Erasing Kinetics of non-quencher bearing Eraser strand.....	63
Figure 4.8: Erasing Displacement Reactions using 4-way branch migration in a 2-strand probe complex.....	65
Figure 4.9: Labeling and removal of Cy5 fluorophores from protein markers using the 4-way-2-strand probe complex.....	67
Figure 5.1: Labeling and erasing reactions using a 4 way-3-strand probe	73
Figure 5.2: Scheme showing the self-assembly of leucine zipper -protein LG with ZR-ELS6-DNA and antibodies.	75
Figure 5.3: Comparison of immunofluorescence image of labeled MT	76
Figure 5.4: Hybridization chain reaction mechanism	78
Figure 5.5: Native PAGE-gel showing the selective amplification and de-amplification (L4) of 3-strand probe construct via the HCR method.....	79

Chapter 1

1.1. Overview

The diagnosis, staging and clinical management of cancer and other diseases is becoming increasingly reliant on the identification and quantification of phenotypic indicators, called 'biomarkers', in histological samples. However, immunohistological methods remain substantially restricted by the fact that only a few biomarkers can be examined on a single biological sample. This is because many fluorescent dyes possess significant overlap in their excitation and emission spectra, and consequently cannot be completely distinguished using standard fluorescence microscopy methods. Furthermore, the immunoreactions used to label protein biomarkers are typically irreversible, prohibiting the removal or exchange of fluorescent reporters within a sample. The use of more than one specimen can circumvent this problem in some cases, but whenever heterogeneity across the specimens from a tissue confounds its characterization or is of central importance, diagnostic information is lost. In addition, vital information that can be derived from the spatial distributions of molecular marker levels in histological samples is lost when multiple specimens are used. This key deficiency compromises many analyses, and often prevents one from extracting the maximum amount of

information from a clinical biopsy, particularly when samples are precious and their sizes are limited.

The goal of this project was to surmount this problem by developing and characterizing an erasable, multi-color imaging technology capable of detecting large numbers of molecular markers on a single biological sample. The technology utilizes dynamic nucleic acid complexes as probes for the selective coupling and removal of fluorophores to and from biomarkers via strand displacement mechanism (i.e. the selective exchange of single oligonucleotide strands between complexes of DNA). The system consists of (1) 'targets', which are ssDNA or partially hybridized oligonucleotide strands conjugated to a protein of interest for biomarker recognition in cells, and (2) multi-strand, fluorophore-containing DNA 'probe complexes' that react with the DNA portion of the target in a sequence dependent fashion via a toehold mediated strand displacement reaction to create fluorescent reporting complexes. The removal of the fluorophore (erasing step) follows via the addition of a quencher-bearing single or partially hybridized DNA complex that displaces the target's DNA strand to effectively remove the dye from the marker so that the sample can be reused to detect other markers with minimal interference from prior rounds of labeling. By combining dynamic DNA complexes that operate independent of one another with spectrally-separated dyes, multiple and unique target/probe pairs can be created that associate specifically and can be imaged in parallel.

In order to successfully employ this technology for reiterative, multiplexed molecular biomarker imaging, the design must meet the following criteria; high specificity of targets to respective probe complexes; efficient labeling and erasing to

ensure that fluorescent signals can be used to fully quantify target abundance without the interference of signals from previous rounds of labeling, and short reaction times to allow for multiple rounds of processing on the same sample without loss of sample integrity. To meet these goals three classes of probes were designed and their structure-function relationships elucidated to determine the contributions of complex size, free energy differences between states, and strand displacement on labeling and erasing kinetics and efficiencies on cells.

1.2. Chapter Summaries

This dissertation describes the development and characterization of an erasable multi-color imaging technology capable of detecting large numbers (several tens and potentially hundreds) of molecular markers on a single sample termed Multiplexed (multi-color) and Reiterative Molecular Biomarker Imaging (MRMBI). This technology utilizes dynamic nucleic acid complexes (i.e. DNA complexes that take advantage of non-equilibrium dynamics of DNA hybridization) as probes to detect molecular markers (biomarkers) on cells and tissues via the selective coupling and removal of fluorophores to and from these markers. It demonstrates that dynamic DNA probes can facilitate multiplexed and reiterative fluorescent labeling reactions on specific intracellular proteins, and should increase the diagnostic power of individual histological examinations thereby opening the door to more sophisticated analyses of cell phenotype and its functional relationship to disease.

Chapter 2 provides background and motivation for this work and briefly describes the importance of immunohistochemical/immunocytochemical techniques to disease diagnostics and prognosis as well as the need for erasable multiplexed fluorescence based immuno-based analyses. Additionally, this chapter provides an overview and the limitations of the current technologies available for in situ multiplexed marker imaging on a single sample. The chapter concludes with a summary of MRMBI technology and its mechanism of operation.

Chapter 3 details the design principles behind the use of dynamic nucleic acid complexes for Multiplexed and Reiterative Molecular Biomarker Imaging. It describes the design and characterization of several dynamic nucleic acid complexes as MRMBI probes in solution and on immobilized surfaces using a microarray platform. The selectivity and effectiveness of these probes in labeling their respective immobilized markers is also evaluated. In addition, the feasibility and utility of these probes for erasable and multiplexed imaging is assessed. Finally, the use of dynamic DNA complexes as imaging probes to reiteratively image targeted molecular markers in fixed cells is demonstrated.

Chapter 4 provides an assessment and comparison of three different classes of imaging probes (probe constructs) to elucidate the contributions of complex size, free energy differences between states and strand displacement reactions on the labeling and erasing efficiencies on fixed cells. The kinetic rates of reaction of the three classes of imaging probes are compared to their calculated net free energy of reaction to determine which factor results in a more efficient and effective probe design. The chapter concludes with the development of an optimized dynamic nucleic acid probe system based on the

results obtained from the characterization of the three probe constructs with short reaction times, non-prohibitive cost of production, highly efficient labeling and effective erasing steps.

Chapter 5, the concluding chapter, provides a summary of the development and characterization of dynamic nucleic acid complexes as imaging probes for erasable multicolor imaging of molecular biomarkers in cells. The implications of the results obtained from this work and its applicability is discussed. The chapter concludes by providing a brief summary of results from on-going experiments that give new insights into the applicability of MRMBI to expanding the utility of immunofluorescence imaging.

Chapter 2

Background

This chapter provides background and motivation for this work and briefly describes the importance of immunohistochemical/immunocytochemical techniques to disease diagnostics and prognosis as well as the need for erasable multiplexed fluorescence based immuno-based analyses. Additionally, this chapter provides an overview and the limitations of the current technologies available for in situ multiplexed marker imaging on a single sample. The chapter concludes with a summary of Multiplexed and Reiterative Molecular Biomarker Imaging technology and its mechanism of operation.

2.1. Motivation

In recent years, tremendous strides have been made in the molecular characterization of malignant and non-malignant diseases in human and the signaling pathways underlying their evolution and progression.¹⁻² A variety of technologies now make it

possible to profile molecular alterations associated with these conditions, including those that assess changes in genetic features (e.g., DNA copy number, chromosome translocations, mutations), epigenetic changes (e.g., chromatin methylation, acetylation, imprinting), gene expression changes, alternative transcripts, regulatory molecules (e.g., microRNAs, non-coding RNAs), protein changes (levels, posttranslational modifications, cellular localization, interaction partners), as well as metabolic states²⁻⁶. The identification of disease-associated changes through molecular and cellular analyses in tissues can be vital for cancer diagnosis, prognosis, and for monitoring patient responses to various treatments (i.e., intermediate, surrogate endpoints) during clinical trials,³⁻⁴ thereby providing opportunities for the personalization and clinical management of cancer.⁵⁻⁷

While there are increasing efforts to develop methodologies for measuring molecular markers through non-invasive means (e.g., serum analysis, surface tissue swabs, molecular imaging)⁸⁻⁹ and through sensitive, high-throughput technologies (e.g., protein lysate arrays,¹⁰⁻¹¹ phosphorylation chromatography¹²), immunohistological analysis remains an invaluable technique that allows spatially-dependent changes in molecular marker levels within cells and tissues to be detected.

Immunohistological analysis offers advantages over bulk analyses, since it allows the spatially dependent expression patterns of RNA and proteins to be delineated in cells and tissues. Several antibody and nucleic acid-based fluorescent probe technologies have been developed for marker imaging and the sensitivity of these probes has been continually improved¹³⁻¹⁵. Yet, many contemporary cytological studies now require increasingly comprehensive molecular pathway analyses to characterize the network-

level properties of cells and resolve functional relationships between cell phenotypes and their tissue distribution¹⁶⁻¹⁸. In such cases, the number of markers one seeks to examine can easily exceed the number of probes that can be used simultaneously for detection due to the spectral overlap of the reporting dye molecules. Thus, various biological studies stand to benefit from methods that allow a greater number of molecular markers to be visualized on the same tissue sample so that any correlations between their distributions might also be analyzed, but immunohistology remains substantially limited in this respect.

2.2. The need for multiplexed immuno-based imaging techniques

Immunofluorescence is a technique that exploits the specificity of antibody-antigen binding to detect specific molecular markers in cells and tissues, providing clinically relevant information for early detection, diagnosis, prognosis, and assessment of responsiveness to intervention. Because Immunofluorescence requires fluorophores for the visualization of markers, the spectral overlap of these fluorophores and the ability to separate them well from one another dictates the number of markers that can be analyzed on the same sample. Although there are several dyes whose peak excitations and emissions span from the UV (~300nm) to the far red (~800nm), the breadth of these spectra (~200 – 300nm) generally permits evaluation of only three markers per sample. Spectral de-convolution techniques make it possible to increase this number (7 to 10 markers can be detected within the cells or tissues),¹⁹⁻²¹ but these capabilities are still insufficient for many applications. For such applications, multiple samples can be used to

image different molecular markers using the same dyes,²² but unfortunately many tissue biopsies are small, limiting the number of sample tissue sections that can be obtained.

For example, in a recent chemoprevention trial at M. D. Anderson Cancer Center involving the use of erlotinib (a tyrosine kinase inhibitor targeting EGFR) to prevent oral cancer development in high risk individuals,²³ our collaborator Dr. Hittelman and fellow researchers were interested in examining multiple molecular markers (e.g., EGFR status, ERK status, cyclin D1 isoforms, proliferation markers, E-cadherin, vimentin, pSrc, TGF α , caspases, FHIT, hTERT, STAT3 status, pAKT, COX2, etc) associated with EGFR pathway. Even though comprehensive analyses of all of these markers would have been helpful for determining cancer risk and the impact of EGFR targeting on premalignant oral tissue, the actual markers that were evaluated had to be cut down and prioritized because the oral biopsies that were available were small.

Many types of studies are restricted substantially by the need for parallel tissue sections for detection of multiple molecular markers, particularly those that rely on analyses of different markers within the same cells. For example, the phosphorylation of proteins at different amino acid positions can reflect their activation state, and relationships between these post-translational modifications and spatially-dependent expression patterns of proteins within cells and tissues can have important pathophysiologic implications.²⁴ Optimally, these examinations would all be made on the same cells so that marker distribution variations between cells would not cloud the assessment. Other types of analyses are simply not possible when multiple samples are examined, especially those that focus on examining rare cells such as stem cells and tumor initiating cells also known as cancer stem cells (Figure 2.1)²¹. The identification

of such rare cells in the tissue implicitly requires the simultaneous use of several markers because of heterogeneity in tissue sections and in the case of stem cells (cancerous or not), information about the spatial localization of these multiple markers are important. Without multiplexed imaging technologies, it is not possible to expand the set of markers assayed, and our ability to characterize these rare cells will remain limited. The current approaches used in identifying tumor initiating cells (cancer stem cells) are discussed below, as well as how advances in multiplexed imaging technologies will facilitate their detection.

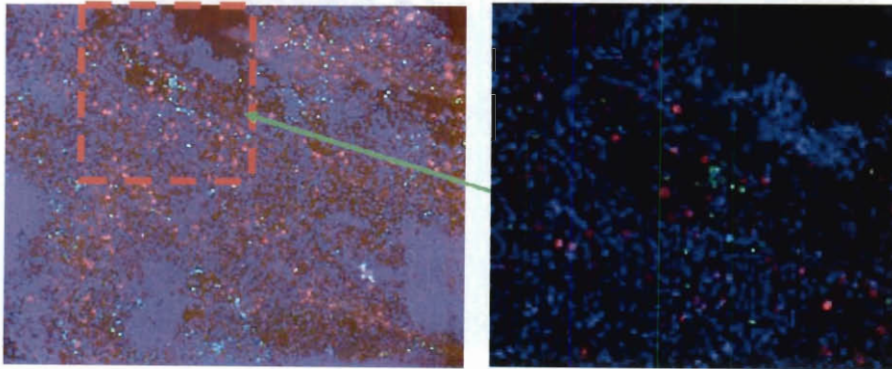


Figure 2.1 : Tissue section of a H460 lung tumor xenograft²¹.

Tumor bearing animals were pulse treated with IdUrd (green label) 10 days earlier to mark the DNA of replicating tumor cells and with ClrdUrd (red label) just prior to tumor harvest. The slides were counterstained with DAPI to mark the nuclei. The cells with green stain are the long label retaining cells and are believed to include the dormant “cancer stem cells” that stopped proliferating early during tumor development. The red cells are the progenitor cells that continue to proliferate as the tumor enlarges

2.2.1. Example Application: Cancer Stem Cell Identification

Multi-color immunohistological analysis is essential in the identification of rare cells such as tumor initiating cells (or cancer stem cells), which can constitute about 0.1 to 1% of a cancerous tumor,²⁵ are suggested to be resistant to radiation therapy and have been implicated in the recurrence of cancer after radio-chemotherapy²⁶⁻²⁷. Identification of these cells not only increases our understanding of the biology of cancers, it also aids in the development of therapeutic strategies that target these cells thereby improving treatment outcomes.

In contrast to many other cell types²⁴, there is currently no single biomarker available that can be used to detect these cells; rather, their identification relies on the presence and/or absence of a combination of multiple cell surface markers,²⁵ signaling pathway indicators and stem cell niche adhesion molecules involved in stem cell activation and mobilizations²⁷. The need to use these markers in combination is driven, in part, by the fact that some of them are prevalent in normal cells, making their individual diagnostic power weak. For example, in the detection of breast cancer tumor initiating cells, Al Hajj et al²⁵. labeled suspensions of tumor cells with antibodies specific for the lineage markers CD2, CD3, CD10, CD16, CD18, CD31, CD64 and CD140b and used fluorescence activated cell sorting (FACS) to separate lineage positive cells from lineage negative cells.

The cells were further probed with the cell surface markers ESA, CD44 and CD24. Cells with various combinations of these markers were isolated by flow cytometry

and implanted into NOD/SCID mice fat pad. Al Hajj et al²⁵ found that cells that were lineage negative, CD24 negative but ESA and CD44 positive exhibited the stem-like ability to form tumors that contained heterogeneous mixture of non-tumorigenic and tumorigenic cells similar to those in the original tumor (Figure 2.12).

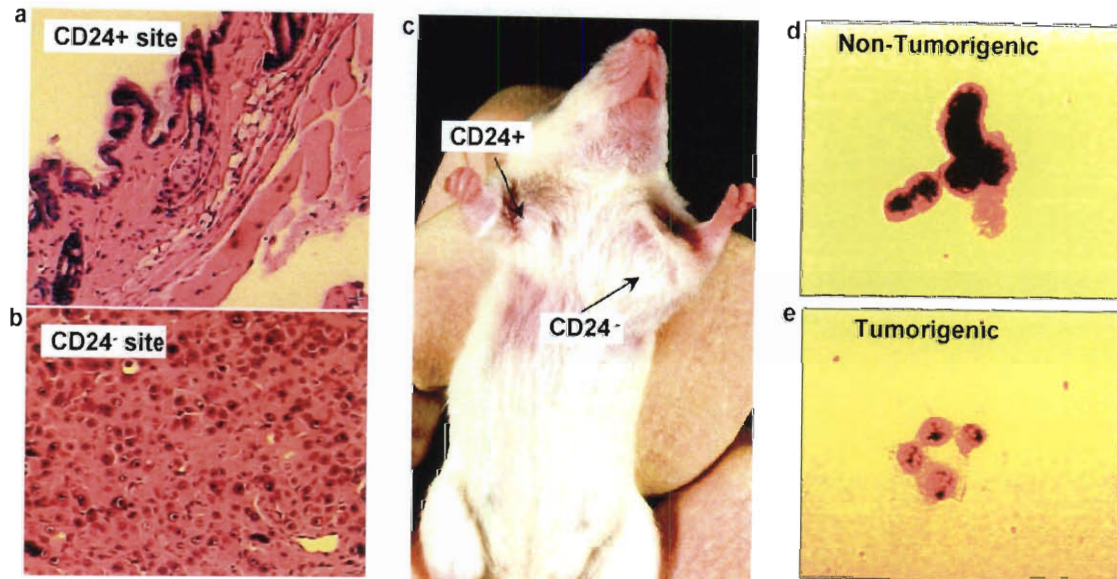


Figure 2.2: Identification of tumorigenic breast cancer cells.

Histology from the CD24+ injection site (a; $\times 20$ objective magnification) revealed only normal mouse tissue, whereas the CD24-/low injection site (b; $\times 40$ objective magnification) contained malignant cells. (c) A representative tumor in a mouse at the CD44+CD24-/lowLineage- injection site, but not at the CD44+CD24+Lineage- injection site. T3 cells were stained with Papanicolaou stain and examined microscopically ($\times 100$ objective). Both the nontumorigenic (d) and tumorigenic (e) populations contained cells with a neoplastic appearance, with large nuclei and prominent nucleoli.

The flow cytometer is inadequate for use in clinical settings where biopsied samples are used because of the need for large sample sizes in analyses. In addition, only cells in suspension can be used in flow cytometry, a criteria that results in loss of spatial information critical in ascertaining niche contributions.

In view of the fact that, the identification of cancer stem cells in tissue implicitly requires the simultaneous use of several markers, the availability of technologies that allow the successive and/or direct detection of a plethora of markers in a single clinical sample will greatly facilitate the understanding of cancer stem cells and also help to provide more targeted therapies.²⁵

2.3. Current Technologies

A number of significant technical issues must be addressed to perform multiplexed marker analyses on a single specimen²⁸. There are two main approaches to addressing this limitation: the first focuses on improving microscope optics and developing new dyes to increase the number of spectrally-separated imaging channels that can be used simultaneously; the second approach is to develop a technique that would allow the serial use of the same “channel” for multiple markers. Both approaches face unique challenges to their application to multiplexed marker identification via immuno-based methods.

2.3.1. Development of optics and dyes for multiplexed marker imaging

In recent years, numerous organic and inorganic dyes have been developed that span a good proportion of the electromagnetic spectrum (from UV to IR),²⁹⁻³² providing a plethora of spectrally distinct dyes available for multicolor imaging. However, these dyes frequently have broad emission spectra resulting in signal spill over (crosstalk) with dyes that have overlapping emission spectra. Quantum dots (QD) outperform single-molecule dyes because of their intrinsically high quantum efficiencies, photostabilities³²⁻³⁴, and narrow emission spectra. For immuno-based detection, antibodies are conjugated to the surface of a quantum dot prior to being used to label molecular markers. The conjugation of antibodies on quantum dot surface, results in an increase in size of the quantum dot as well as an unorganized distribution of antibodies on QD's. Thus, despite its numerous attributes, QD-based labeling approaches in many clinical applications³⁴ have had limited success in part due to the large size of QDs (compared to antibodies) and the unorganized presentation of antibodies on their surface, which affects the efficiency and selectivity of antibodies. Regardless, both QD's and organic/inorganic dyes need to be spectrally separated with optical filters to ensure minimal crosstalk between channels, which limits the number of colors that can be simultaneously probed to about 3 or 4. This is because optical filters that can adequately separate the excitation and emission spectra of fluorophores with overlapping spectra may decrease the signal strength or eliminate signal from one fluorophore because of its spectral bandwidth.

This issue is being addressed through the development of hyperspectral imaging microscopes that have the ability to de-convolve spectrally overlapped dyes into their unique components when used in imaging applications, and have been used to detect up

to 10 different markers on a single sample³⁵. These instruments are often expensive and not widely available, so alternative approaches that do not rely heavily on optics must be developed to complement advances in microscope technologies. The number of markers that can be evaluated on an individual biological sample could be increased if it were possible to remove fluorescent probes from cells such that new markers could be labeled and detected using the same fluorescent reporting molecules. Sequential analysis of biological specimens takes advantage of current fluorophores and epifluorescent microscopes to detect multiple markers on a singular sample by re-probing the same sample multiple times. Sequential analysis when combined with hyperspectral imaging microscopes has the potential to greatly increase the number of markers evaluated on a single biological sample.

2.3.2. Sequential analysis of biological specimen for multiplexed marker imaging

Re-labeling of a sample can be accomplished in two ways, either by removing the antibody (and the dye in doing so), or by removing the dye alone. Sequential removal of antibody probes has been accomplished by treating samples with caustic chemicals prior to the application of the next set of antibody probes³⁶⁻³⁷. Alternatively, high temperatures have been used to inactivate antibodies prior to the application of the next set of antibody probes³⁸. These harsh methods have limited utility because they tend to compromise both the molecular-level epitopes targeted by antibodies and the larger scale morphological features of cells and tissues.

One method of “removing” dyes involves iterative re-staining where the same fluorophore is used to label and re-label different antigens on the same cell. Instead of physically removing the dyes, their signals are accounted for at each round of imaging

and removed from subsequent rounds during data processing. A laser scanning cytometer is required to track intensities from current and previous iterations of staining and to determine the fluorescence shifts from the subsequent labeling⁴⁰. Due to the addition of dyes through sequential labeling, detection of low level signals is difficult; saturation of channels is a problem, and the addition of the same dye in subsequent imaging steps results in high background noise.

An alternate method involves the bleaching of the fluorophores following each sequential labeling⁴⁰. This could be done either through differential photobleaching, (i.e. identifying different fluorophores based on their unique photostability signatures)⁴⁰ or by photo-destruction of fluorophores. Both approaches involve exposing the sample to laser irradiation for 10 mins or more, which can adversely affect sample integrity when performed multiple times. In addition, bleaching does not completely remove the signal from prior labeling, resulting in residual accumulation in subsequent processing steps that compromises analyses. Therefore, new approaches are needed for multiplexed molecular marker analyses of cells whereby the same sample can be probed and re-probed for multiple markers sequentially.

2.4. The solution: Multiplexed and Reiterative Molecular Biomarker Imaging (MRMBI)

Recent advances in the field of DNA nanotechnology have facilitated the creation of various dynamic DNA and RNA complexes that can function effectively as programmable logic gates⁴¹⁻⁴³, chemical amplifiers⁴⁴⁻⁴⁵, and reconfigurable molecular structures⁴⁶. A key feature of these complexes is that, instead of classical hybridization reactions, they can operate via a process called strand displacement⁴⁷ – the exchange oligonucleotides possessing partially or fully identical sequences between different thermodynamically-stable multi-strand complexes (examples are shown in Figures 1 and 2). Using this mechanism, long nucleic acid complexes possessing many matched base pairs can be hybridized and de-hybridized multiple times at room temperature. Since strand displacement reactions are sequence dependent, and tend to be more sensitive to base mismatches than classical hybridization reactions⁴⁸, different dynamic complexes can be designed to operate independently of one another, or, alternatively, integrated into reaction networks that can perform complex computations that are programmed in oligonucleotide sequence⁴⁹⁻⁵⁰. Such capabilities now offer opportunities to create new classes of molecular probes for molecular-cell analyses⁵¹⁻⁵².

The unique potential of dynamic nucleic acid complexes for molecular-cell analyses are beginning to be realized. Engineered RNA hairpin devices have been used as ‘smart’ therapeutic technologies that can both detect genetic disease indicators *in vitro* and, in response, produce double stranded oligonucleotide polymers that selectively trigger cell apoptosis⁵³. Multiplexed (5-color) *in situ* detection of mRNA transcripts in fixed drosophila embryos has also been demonstrated⁵².

The goal of this project is to develop utilize dynamic nucleic acid complexes in the development of an erasable multicolor molecular marker analysis technology (termed multiplexed and reiterative molecular biomarker imaging) capable of detecting large numbers (several tens and potentially hundreds) of molecular markers on a single clinical sample. The system consists of (1) ‘targets’, which are single or partially hybridized DNA strands conjugated to a protein of interest for biomarker recognition in cells, and (2) multi-strand, fluorophore-containing DNA ‘probe complexes’ that react with the DNA portion of the target in a sequence dependent fashion to create chemically switchable fluorescent reporting complexes. The addition of a quencher-bearing ssDNA displaces the target’s DNA strand to effectively remove the dye from the marker so that the sample can be re-imaged for other markers with minimal interference from prior rounds of labeling. The ability to control the switching of these complexes is made possible by the use a process known as strand displacement. This is the selective exchange of single or partially hybridized oligonucleotide strands possessing fully or partially identical sequences between different thermodynamically stable multi-strand complexes.⁴⁷

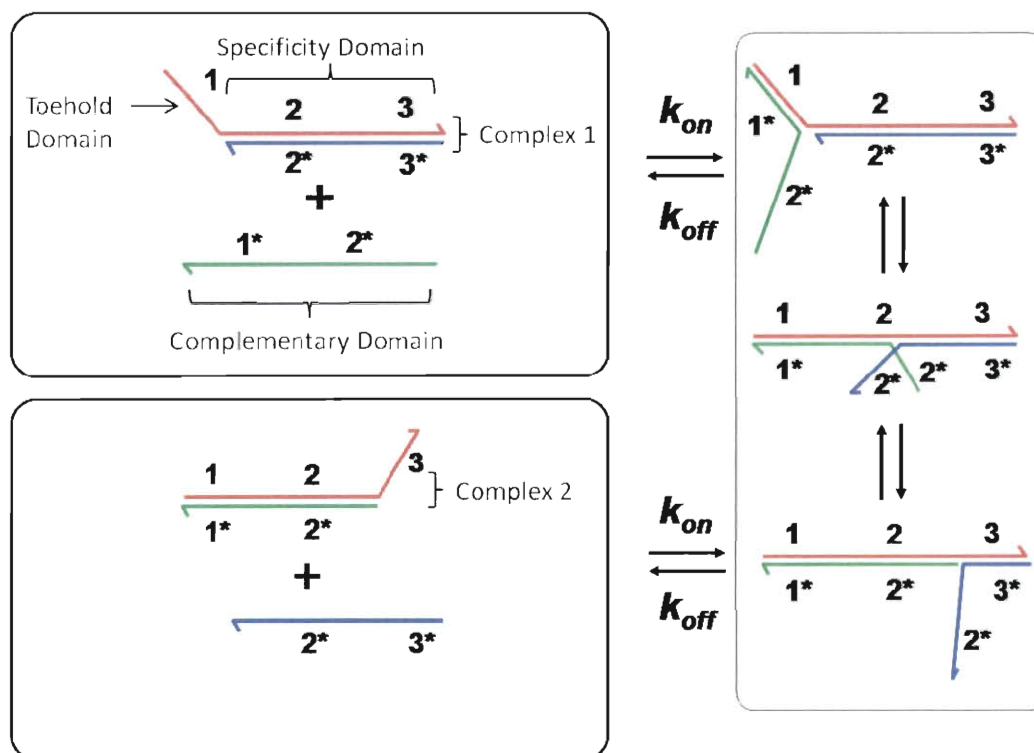


Figure 2.3: DNA strand displacement mechanism.

Domains 1, 2 and 3 are complementary to domains 1*, 2* and 3* respectively. Domain 1 is a toehold domain; Domains 2 and 3 are specificity domains (Top left); Domain 3 is a toehold domain; Domains 1 and 2 are specificity domains (Bottom left); Step by step illustration of strand displacement (Right).

Figure 2.3 illustrates the strand displacement mechanism. Here domains 1, 2 and 3 are complementary to domains 1*, 2* and 3*. Domains 2 and 3 are specificity domains because any base mismatch within their sequence will slow down the strand displacement reaction. The strand displacement reaction proceeds when the single stranded portion of complex 1 (domain 1) serves as a toehold through which domain 1* of the green strand (complementary domain) can latch on to and initiate the invasion of complex 1 and subsequently displace the

blue strand in a sequence specific manner. This creates a new partially hybridized complex (complex 2) with an un-hybridized domain 3 which serves as a toehold that can facilitate another strand displacement reaction.

Thus, the strand displacement mechanism allows DNA complexes to go through multiple states by being hybridized and un-hybridized numerous times at room temperature before equilibrium is reached. Strand displacement reactions are sequence dependent and more sensitive to base mismatches than classical hybridizations reactions⁴⁸ thus several dynamic DNA complexes can be designed to operate independent of one another and used to image different molecular markers simultaneously.

By combining dynamic DNA complexes with self assembled DNA-conjugated antibodies, the project develops a new technology that utilizes the immunoreactivity of antibodies for antigen recognition and the strand displacement capabilities of dynamic DNA complexes for fluorophore exchange. This capability allows the same color fluorophores to be used reiteratively, via multiple rounds of fluorescent microscopy, to detect different sets of protein markers as illustrated in Figure 2.4. Subsequently, much more comprehensive analyses of molecular marker levels can be performed since the spectral overlap of fluorescent dyes will no longer limit the number of molecular markers that can be visualized on individual samples using fluorescence microscopy. For this application, the strand displacement reactions driving the labeling and erasing of markers must be selective and efficient (high labeling and dye removal yields) to facilitate quantitative assessments of markers levels and to ensure residual signals left on a sample after an erasing step do not compromise subsequent markers analyses.

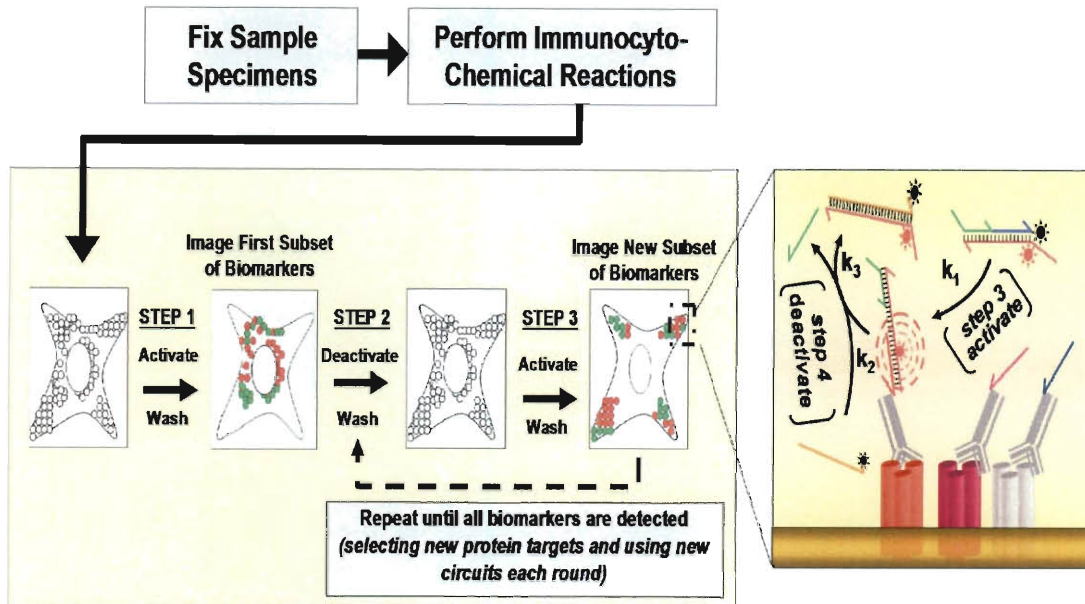


Figure 2.4: Multiplexed biomarker imaging via Multiplexed and Reiterative Molecular Biomarker Imaging.

2.4.1. Adaptation of Entropy-driven DNA circuits for Multiplexed and Reiterative Molecular Biomarker Imaging

The initial imaging probes were designed around a class of artificial biochemical circuits developed by Zhang *et al*⁵⁴. These circuits utilize simple DNA complexes as both designable catalysts and as substrates in biochemical reactions that involve the selective exchange and release of specific single-strands of DNA between reaction components. Individual reactions that make up these circuits are driven forward by configurational entropy: the number of components present in solution is increased as these reactions proceed to completion.

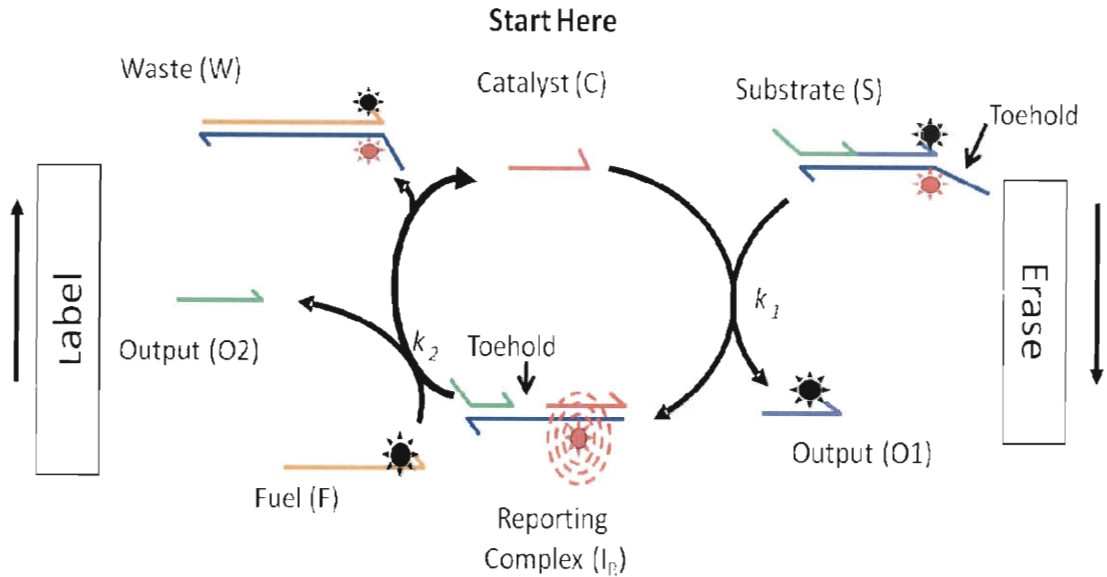


Figure 2.5: The adaptation of entropically driven circuit for MRMBI.

This mechanism is optimal for biomarker imaging since these circuit designs do not depend on enthalpy-driven reactions to function (e.g., via a net increase in base-pair formation or conformational changes of engineered DNA complexes). These entropy-driven reactions function robustly at ambient temperatures and in mild buffers (e.g., 20 mM Tris, pH 7-8). Consequently, this technology will facilitate fluorophore reutilization while minimizing perturbations to the morphology and composition of cells and tissues.

The artificial biochemical circuit adapted for multiplexed and reiterative molecular biomarker imaging is illustrated in figure 2.5. Adopting the nomenclature of Zhang *et al*⁵⁴, the entropically driven circuits are composed of three main components: a single stranded catalyst strand (C), a partially hybridized three-strand DNA complex called the substrate (S), and a ssDNA fuel strand (F). DNA circuit reactions are initiated by the

binding of the catalyst strand to a 6 bp ssDNA domain at one end of the substrate called a 'toe-hold' domain. The binding of the catalyst to the toehold domain results in the invasion of the substrate complex by the catalyst strand leading to the displacement of the 'output strand' (O1) within the substrate complex by the catalyst strand. This strand displacement reaction produces an intermediate state complex (I_R) that now possesses a new, internal toe-hold composed of 4 unmatched base pairs. The circuit reaction cycle can then be completed via a second strand exchange reaction that is initiated by the binding of F to the toe-hold in I_R . This reaction releases both the catalyst and a second substrate strand (O2), and produces a single linear duplex called the waste complex (W). Thus the catalyst can engage in successive strand displacement reactions in the presence of a substrate complex to produce the intermediate reporting complex.

To convert circuits into imaging probes, fluorescent dyes (Cy3 or Cy5) and quenchers (Iowa Black) were incorporated into the substrate complexes. These molecules are positioned within the substrate complex such that the displacement of output O1 by the catalyst strand produces an I_R complex (reporting complex) that contains an unquenched fluorophore – which is the labeling reaction. The addition of a quencher-bearing fuel strand can then be used to selectively remove the fluorophore-bearing strands from the catalyst, resulting in a quenched duplexed waste complex – which is the erasing reaction. For biomarker recognition, the ssDNA catalyst strand is conjugated to a protein (ex. Antibody) that binds to biomarker of interest thus enabling its detection during the labeling step. In addition, the technology is amenable to the detection of a plethora of biomarkers so long as the catalyst can be conjugated to their respective ligands (antibody, mRNA etc).

In conclusion, multiplexed and reiterative molecular biomarker imaging opens up the possibility to analyze the genomic and proteomic properties of whole cell populations on the single cell level. It also facilitates our understanding of complex processes in health and disease and could be an important tool for predictive and preventative medicine.

Chapter 3

Multiplexed and Reiterative Fluorescence Labeling via DNA Circuitry¹²

3.1. Introduction

In this chapter we show that the dynamic DNA complexes designed by Zhang *et al*⁵⁴ for use as an artificial biochemical circuit can be adapted to create erasable molecular imaging probes that can function at ambient temperatures and in mild, non-denaturing buffers (e.g., Tris-based buffers). Several dynamic DNA complexes are designed and characterized in solution and on immobilized surfaces using a microarray platform. The selectivity and effectiveness of these probes in labeling their respective immobilized markers is also evaluated. In addition, the feasibility and utility of these probes for

¹ This chapter has been published in the following journal article: Duose D. Y, Schweller R. M, Hittelman W. N and Diehl M. R, (2010) “*Multiplexed and Reiterative Fluorescence Labeling via DNA Circuitry*”, *Bioconjug. Chem.* 21(12), 2327-2331.

erasable and multiplexed imaging is assessed. Finally, the use of dynamic DNA complexes as imaging probes to reiteratively image targeted molecular markers in fixed cells is demonstrated.

3.2. Materials and Methods

3.2.1. DNA Circuit Design and Characterization

Two of the dynamic DNA complexes (circuit 1 and 2) used in our experiments was taken directly from Zhang *et al*⁵⁴, the others were designed in-house. The strands for each catalyst, substrate and fuel (eraser) set were designed on a domain basis. The domains were designed according to their functionality. Toehold domains were designed to be short (4-10nt) as this accelerated the initiation of strand displacement reactions. Specificity domains were designed to be thermally stable, and to function in a sequence specific manner. Complementary domains were designed to be complementary only to their respective specificity domains. First, random sequences with 40-50% GC content were generated for each domain, checked for secondary structure using mFold⁵⁵ and then altered by hand to remove bases that cause secondary structure. Next, the domains were concatenated together, rechecked for secondary structure and spurious bindings with mFold⁵⁵. This process was repeated till a satisfactory strand was obtained that contained minimum secondary structure.

Fluorophores and quenchers were incorporated into substrate complexes as shown in Appendix A. Strands outfitted with fluorophores or quenchers were purchased HPLC-

purified. DNA complexes were formed via a thermal annealing procedure: strands were mixed together at a 1:1 stoichiometry in TAE/Mg²⁺ buffer at a final concentration of 3 μ M. The temperature of this solution was then raised to 95 °C and reduced to 25 °C over 90 min. DNA complex formation was verified by 12% nondenaturing PAGE gel analyses using SYBR-Gold staining (Invitrogen).

3.2.2. Conjugation of ssDNA catalyst with artificial protein (leucine zipper)

Oligonucleotides were purchased from Integrated DNA Technologies (IDT) and their sequences are given in Appendix A. The recombinant target protein GFP-Z_E, was designed with C-terminal leucine zipper (Z_E) and 6 \times Histidine tag (6xHis) for purification using standard cloning procedures. Artificial proteins were labeled with catalyst DNA as described in reference 56. Briefly, the C-terminal cysteine of the polymers was reduced and covalently linked to the amine terminated single-stranded DNA catalyst using the hetero-bifunctional crosslinking reagent SMCC⁵⁷ 450 μ L of 10 mg/ml protein in 8M Urea, pH 7.2 was added to 50 μ L of 400 mM TCEP (tris(2-carboxyethyl)phosphine), and the pH was adjusted to 4.5. The C-terminal cysteine in the protein was reduced by incubating the reaction for 1.5 hours at 37°C. While incubating, a Nap-5 column was equilibrated with 10 ml of 8 M Urea. After incubation, 500 μ L of the reduced protein was added to the Nap-5 column to remove excess reducing agent, and eluted with 1 ml 8M Urea pH 7.2, yielding a final volume of 1 ml. Half an hour into the protein reduction reaction, the amine DNA was reacted with the NHS-ester of the sulfo-SMCC by combining 100 μ L of amine-terminated DNA (100 μ M), 100 μ L conjugation buffer (20 mM NaH₂PO₄, 80 mM Na₂HPO₄, 150 mM NaCl and 1 mM EDTA pH to 7.3), and 60 μ L sulfo-SMCC (2 mg sulfo-SMCC/60 μ L DMF). This mixture was allowed

to react at 37°C for 1 hour. A second Nap-5 column was equilibrated with 10 ml 1X PBS. After incubation, the DNA reaction was diluted by adding 240 μ L of conjugation buffer, and was then transferred to the Nap-5 column and eluted with 1 ml 1X PBS, yielding a final volume of 1 ml DNA. The DNA and reduced protein volumes were then added in 1:1 ratio in microcentrifuge tubes, wrapped in foil, and agitated with a shaker for 2 hours at room temperature. The shaker was then moved to 4°C, and allowed to react overnight. The sample was then purified using FPLC with a Hi trap Q XL 5 mL column on a 1-100% NaCl gradient over 12 CV in 20 mM Tris buffer (pH 8.3). The fractions, identified by UV absorbance were collected and analyzed using an SDS-PAGE gel treated with Stains-All to verify the presence of both DNA and protein in the sample. Fractions containing the DNA/protein conjugate were combined, lyophilized, dissolved in TAE+12.5 mM Mg(OAc)₂, aliquoted, and then stored at -20°C.

3.2.3. Microarray Procedures

The DNA microarrays were printed on Vantage silylated aldehyde slides (CEL Associates) using SMP3 pins (ArrayIt) and a custom fabricated microcontact printer. Arrays were fabricated by spotting solutions of 3'-amine labeled C strands (10, 5, 1, 0.5, 0.1, and 0.05 μ M stocks for gradient experiments and 2 μ M stocks for all others) in PBS (pH 6.6) containing 30% glycerol. Afterward, the slides were incubated in a humidity chamber for 6 h. Free aldehyde groups on the slides were then quenched for 5 min in a sodium borohydride solution (3:1 PBS/EtOH 2.5% NaBH₄). Slides were blocked for 2 h in 4 \times SSPE buffer (600 mM NaCl, 40 mM NaH₂PO₄, and 5 mM EDTA) containing 0.1% BSA, washed 3 times with 4 \times SSPE buffer with 0.1% SDS, rinsed with milli-H₂O, and dried under nitrogen.

Microarray labeling/activation and dye removal/deactivation reactions were performed using a static incubation procedure or with a hybridization station (TrayMix2: ArrayIt) that provides active mixing of reagents over the slide. For static incubation, Gene frames (AbGene) were affixed over the arrays to create a reaction chamber containing 5 pmol substrate and OC1 consumption complex in TAE w/Mg²⁺. The arrays were incubated overnight. The Gene frames were then removed and slides were washed 3 times in 4× SSPE buffer, rinsed again in milli-H₂O, and dried under a nitrogen stream. Deactivation and reiterative labeling experiments were performed similarly by affixing new Gene frames to the slides and repeating the incubation procedure. Microarray analysis was performed using *ImageJ* (<http://rsbweb.nih.gov.ezproxy.rice.edu/ij/>).

3.2.4. Cell Culture and Labeling

HeLa cells were cultured in an 8-chambered coverslips (Lab-Tek) for 24 h in DMEM supplemented with 10% FBS, 50 µg/mL penicillin, and 50 µg/mL of streptomycin. For GFP labeling experiments, the media was replaced and the cells were transfected with GFP-Z_E DNA using Fugene (Roche) transfection reagent under the manufacturer's protocol.

To label cells, coverslips were washed once with PBS and fixed with freshly prepared 4% paraformaldehyde (PFA) in PBS pH 7.2 for 30 min. The cells were then washed twice for 2 min with PBS, permeabilized for 5 min with 0.2% Triton X-100, washed twice again with PBS, and stored overnight at 4 °C. Prior to circuit labeling experiments, the cells were washed again with PBS and then incubated for 2 h with a blocking solution (1% BSA, 1 mg/mL denatured Herring sperm DNA, and 0.5 µM polyT

DNA in PBS). For GFP labeling experiments, the cells were also incubated with 400 nM of Z_R-ELS₆-Cat1 in PBS for 2 h. Excess polymer was then removed by washing twice with PBS prior to circuit-based labeling.

Circuit labeling reactions were carried out by incubating the cells for 1.5 h with 100 nM substrate complex and then washing twice with PBS for 2 min. Dye removal was performed similarly using 1 μM fuel. Before imaging experiments, slides were washed twice with PBS and then mounted on a glass slide using rubber cement.

3.2.5. Fluorescence Imaging and Analyses

All images were collected using a Zeiss Axioplan epifluorescence microscope and are contrasted identically in each figure. Correlations between GFP and Cy5-circuit signals and Cy5 signals produced during sequential labeling of cells (ON₁ and ON₂) were analyzed using a custom program written in *Matlab*.

3.3. Results and Discussion

An illustration depicting our use of entropically driven circuits as imaging probes is shown in Figure 3.1. Adopting the nomenclature of Zhang et al.⁵⁴, these circuits are composed of three main components: a single-stranded “catalyst” strand (C), a three-strand DNA complex called the “substrate” (S), and a “fuel” strand (F).

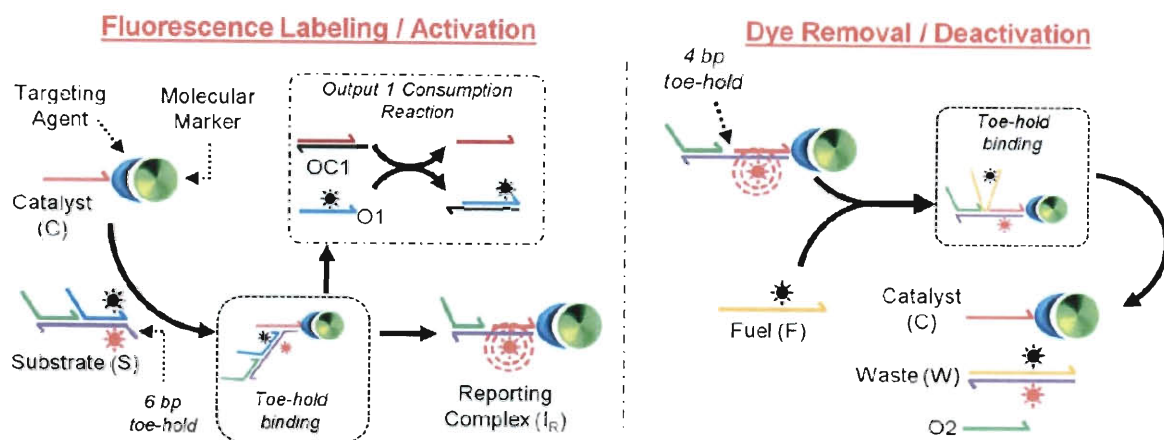


Figure 3.1: DNA circuit-based marker labeling and dye removal reactions.

In the circuit reaction cycle, the binding of C to a 6 nucleotide “toehold” domain at one end of S initiates a strand displacement reaction that releases an “output” strand (O1) and produces an intermediate-state complex (I_R) that possesses a new, internal 4 bp toehold domain. The binding of F to this toehold then initiates a second strand displacement reaction that releases C from the I_R complex and produces a “waste” product (W). To convert these complexes into imaging probes, catalyst strands can be appended to targeting agents that bind to specific molecular markers. The circuit substrates are modified by incorporating fluorescent dye (Cy3 or Cy5) and quencher (Iowa Black) molecules that are positioned such that the reaction of S with C results in an I_R complex that contains an unquenched fluorophore. The dye-bearing strand within I_R can then be removed from the marker and rendered inactive in the waste product by incubation with a modified F strand that carries a second quencher molecule. Overall, the use of quenched substrates and waste products should reduce background fluorescence resulting from potential nonspecific binding of either complex to a sample. Finally, we

note that, in contrast to prior work where C truly functioned as a catalyst⁵⁴, marker labeling and removal in the present application is achieved by performing partial circuit reactions while using C for targeting. We therefore use the term catalyst only for continuity with previous reports.

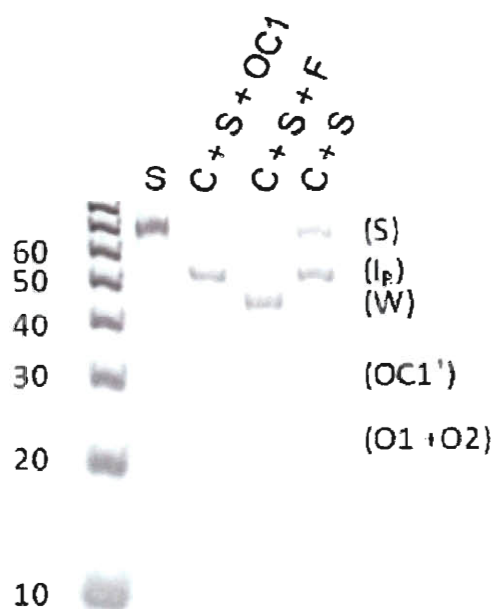


Figure 3.2: Native PAGE gel displaying DNA circuit reaction products

The lanes in the gel can correspond to: lane 1: 10 bp DNA ladder; lane 2: (S); lane 3: partial circuit reaction of S and C that includes output 1 consumption; lane 4: full circuit reaction of S, C, and F; lane 5: partial reaction of S and C that omits output 1 consumption. In all lanes, $[S] = [C] = [F] = [OC] = 200$ nM.

For the labeling and erasing reactions to function efficiently, the formation of the reporting and waste complex respectively must go to completion. To evaluate the efficiency of the circuit-based labeling and dye removal reactions, we first examined distributions of product complexes that were formed upon incubations of S and C, as well as S, C, and F via native PAGE-gel analyses (Figure 3.2). After a partial circuit reaction of S with C, the catalyst strand is bound to the I_R complex through a total of 22 matched base pairs. This complex is therefore stable at room temperature and can be isolated on a gel. Yet, the free energy difference between S and I_R is small ($\Delta G \approx -0.4$ kcal/mol). As a result, the reaction of S and C, when performed using equimolar concentrations, results in an equilibrium distribution of circuit components possessing near-equivalent concentrations of S, I_R , and free, “unlabeled” catalyst (Figure 3.2, lane 5). While this result is not optimal for marker labeling, this reaction can be driven forward by adding a second complex (OC1) that consumes (O1) once it is liberated from S (boxed reaction in Figure 3.1). As demonstrated in other strand displacement systems⁵⁸, the sequestration of O1 shifts the equilibrium distribution of the reaction significantly toward the I_R state (Figure 3.2, lane 3). Alternatively, this distribution can also be shifted using an excess of S relative to C, which should often be the case when labeling catalysts (markers) that are immobilized on a specimen. Nevertheless, the ability to drive strand displacement reactions forward through output sequestration will likely be useful for optimizing dye-labeling kinetics or when local target concentrations within a sample are high.

The removal of dyes from a sample via the reaction of I_R and F constitutes an equally important step in our marker imaging procedure. Here, the use of three-strand S and I_R complexes, as opposed to somewhat simpler two-strand complexes, allows dye-

bearing strands to be displaced from the reporting I_R complex without output sequestration (Figure 3.2, lane 4), since two strands (C and O2) must react simultaneously with final circuit product W for the reverse reaction $W + O_2 + C \rightarrow I_R + F$ to occur. Hence, the dye removal reaction is effectively irreversible.

We next performed a series of DNA microarray experiments that were designed to evaluate the use of DNA circuit complexes as molecular imaging probes (Figure 3.3).

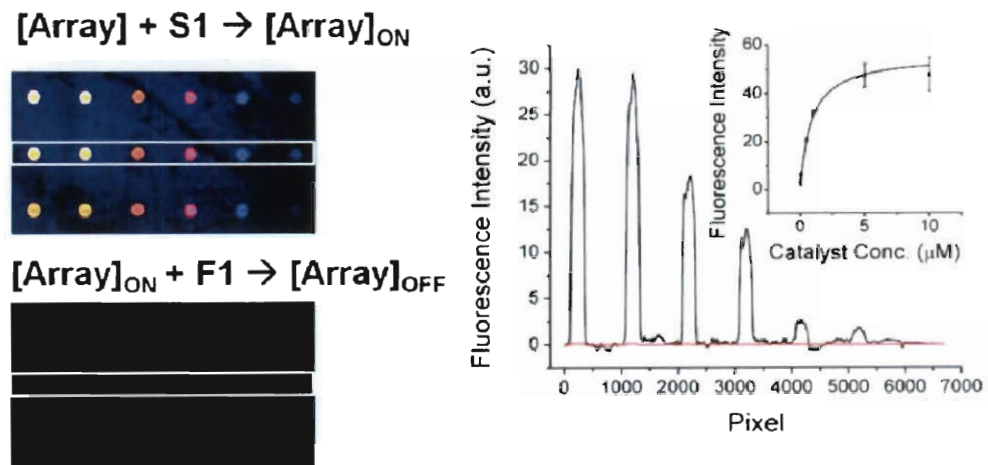


Figure 3.3: Concentration gradient of catalyst microarray.

A catalyst microarray possessing a gradient of spot concentrations (using 10, 5, 1, 0.5, 0.1, and 0.05 μM stock solutions) that was first labelled and then erased via the sequential addition of S (top image) and a fuel strand (bottom image). Images are rendered as heat maps. Plots of the averaged intensity profile of the boxed regions for the labelling (black line) and dye removal (red line) reactions are shown. Average spot intensities can be approximated by the Langmuir equation (inset).

In these experiments, amine-modified C strands are arrayed on the surfaces of glass slides, and the reaction of S with C produces a fluorescent I_R complex that is anchored to the slide surface and can be detected using a fluorescence microarray scanner. Analyses of arrays where a catalyst was printed at variable spot concentrations confirm that the I_R complex of the circuit can function as a reliable reporter of the levels of immobilized catalyst (Figure 3.3). Here, spot intensity profiles can be approximated by the Langmuir adsorption equation, as is commonly found with DNA microarrays⁵⁹⁻⁶⁰. Furthermore, after the same slide is incubated with a fuel strand, each spot disappears and cannot be detected over background autofluorescence signals of the slide (Figure 3.3, bottom left).

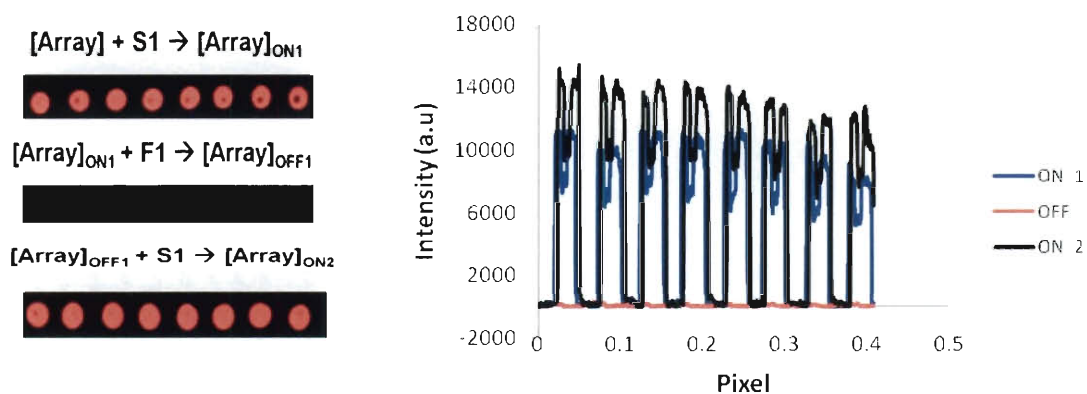


Figure 3.4. Reiterative Labeling of individual arrayed catalysts.

The same catalyst strand was labelled (ON1), erased (OFF) and re-labelled (ON2) on the same slide.

Our microarray experiments also allowed us to demonstrate the use of DNA circuit probes for both multiplexed and reiterative marker labeling. We found that immobilized C strands can be labeled with dyes multiple times via sequential reactions of arrays with S, F, and then a second solution of S complexes (Figure 3.4). Each labeling reaction produced arrays possessing near identical spot intensities.

We also demonstrated that multiple C targets could be labeled and/or erased simultaneously using multiple S and F complexes in a single reaction step (Figure 3.5). In these experiments, five different C strands were printed both as mixtures of two strands and individually on the surface of the slide. The array was then reacted with two different substrate complexes (S1-Cy5 and S2-Cy3), yielding a spot pattern that corresponded directly to the positions of the printed catalysts (C1 and C2). Subsequently, in a single incubation step, this pattern was erased with F1 and F2, and a new spot pattern was generated through a reaction with a second set of substrates (S3-Cy5 and S4-Cy3); spotted lanes where C3 and C4 were printed appear in the scanned image. Throughout this procedure, the fifth printed catalyst strand of the array (C5) remains unlabeled in both scanned images.

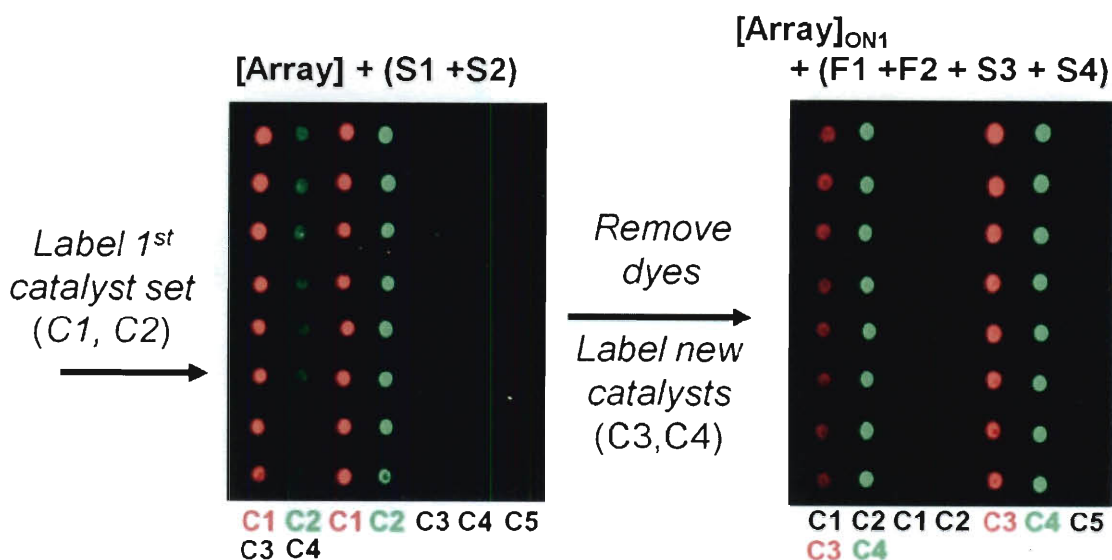


Figure 3.5: Multiplexed and reiterative labeling of multiple DNA catalysts.

The reactions performed on the arrays are indicated above each image.

Thus, these experiments confirm that DNA circuitry can be used for multiplexed and reiterative imaging: two fluorescent dye molecules and two spectral channels of an imaging system are used to detect four distinct markers on the same sample. Importantly, all labeling and removal reactions in these assays were performed using mild processing conditions (room temperature and Tris buffer supplemented with 12.5 mM Mg^{2+}).

We next performed a series of imaging experiments that show DNA circuit complexes can be used to selectively label molecular markers on fixed and permeabilized HeLa cells. Background circuit reactivity was first tested by incubating cell samples with a quenched S complex (100 nM) for 1.5 h. The resulting images show no discernible fluorescence signals and possess signal to background ratios of 1 (Figure 3.6), implying that the substrate complexes have exceptionally low background reactivity with cells. We

attribute this property to the high stability of the duplexed substrate complex, the enhanced sequence specificity of strand exchange over a classical hybridization mechanism⁶¹, and the ability to place dye and quencher molecules in close proximity to one another within the substrate complex.

[Slide] + S1

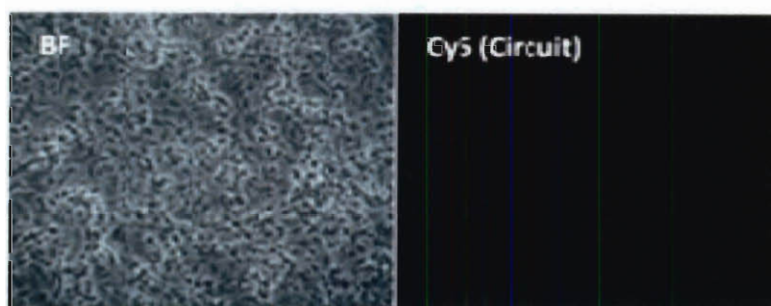


Figure 3.6 Determination of background circuit reactivity

Bright field (BF) and fluorescence images of paraformaldehyde-fixed and permeabilized HeLa cells reacted with a substrate complex that incorporates a Cy5 and a quencher (Iowa Black).

To label markers on cells, we chose to target the DNA circuit complexes to a transfected and expressed green fluorescent protein construct (GFP-Z_E) so that circuit labeling and dye removal efficiencies could be benchmarked directly against an internal standard (

Figure 3.7). The catalyst strand was coupled to the GFP using DNA-conjugated artificial-protein-based polymers (Z_R-ELS₆-ssCat) that we have

previously developed for protein–DNA labeling⁵⁶. These polymers associate with the GFP-Z_E via a heterodimeric leucine zipper complex (Z_E/Z_R: $K_D \sim 10^{-15}$ M). Thus, after incubating GFP-Z_E transfected HeLa cells with Z_R-ELS₆-ssCat and then washing the samples to remove unbound polymer, GFP-Z_E transfected cells can be labeled by a reaction of a circuit substrate complex that carries a Cy5 dye and a quencher.

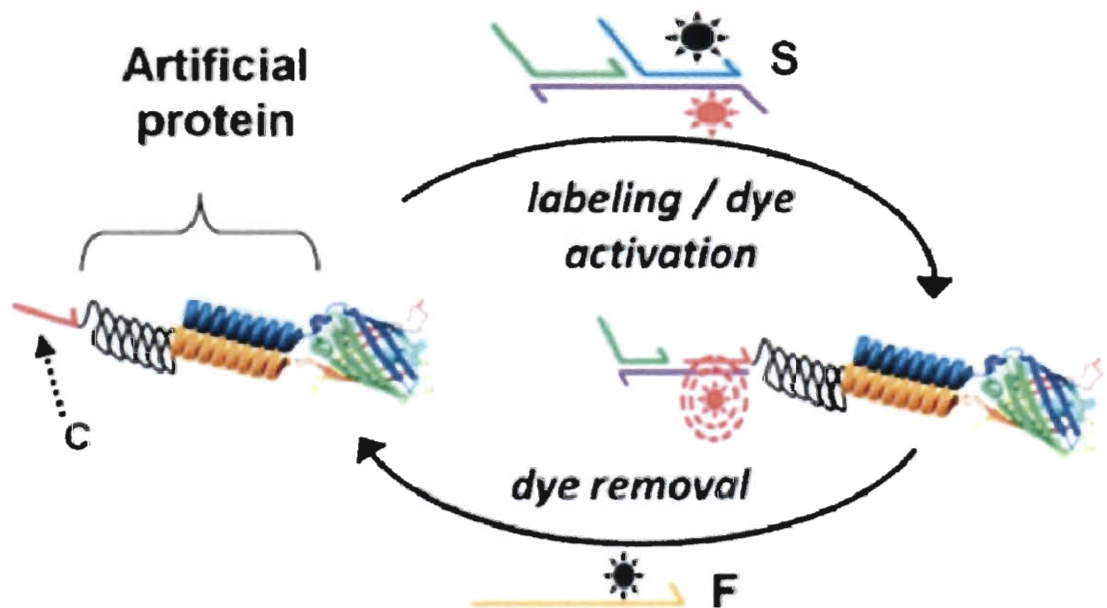


Figure 3.7: Targeting DNA circuit complex to GFP-Z_E.

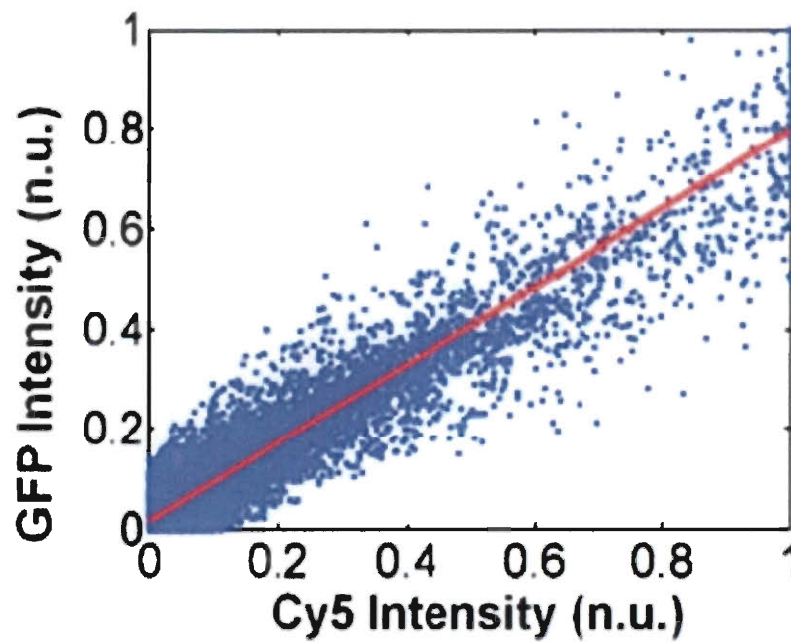
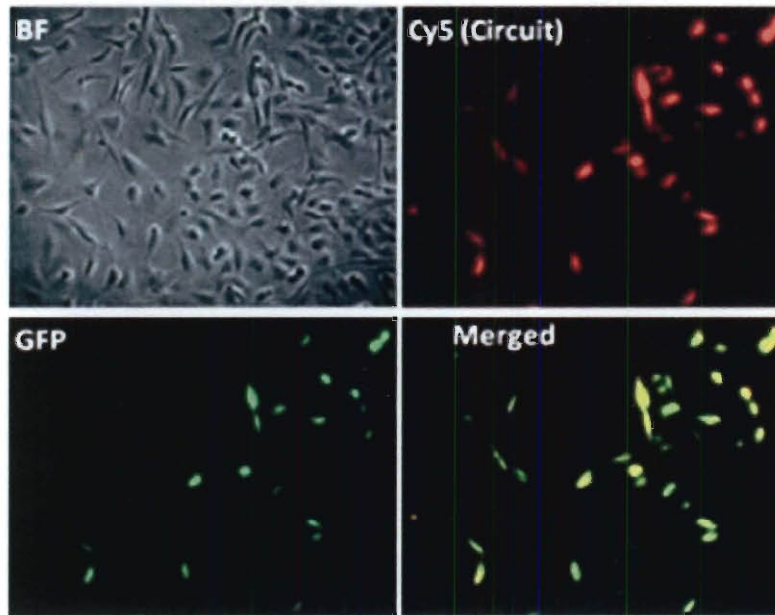
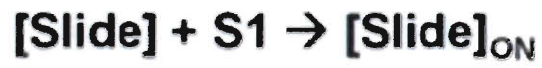


Figure 3.8: Circuit-based labeling of GFP-ZE transfected HeLa cells.

(top) and intensity correlation analyses of GFP-ZE and DNA circuit (Cy5) signals (bottom).

As seen in Figure 3.8, cells that were successfully transfected with GFP-Z_E reacted with S to produce fluorescent signals in the Cy5 channel of the microscope. While cells that were not transfected did not exhibit fluorescence, clear linear correlations are observed between GFP-Z_E and circuit labeling intensities, yielding a correlation coefficient $r = 0.95$.

We also tested whether molecular markers could be labeled multiple times on a single sample of cells without loss of fluorescence signal intensities (Figure 3.8). After a first round of circuit labeling and imaging, fuel strands were added to remove Cy5 dyes from a sample of GFP-Z_E transfected cells. As in our microarray experiments, dye removal reactions are found to be efficient and yield signal to background ratios of 1. In addition, the transfected cells could be labeled and imaged a second time by incubating the sample with a fresh solution of substrate. Bright field imaging showed that a small portion of the cells detached from the slide surface during our manual washing and coverslip mounting procedures. Nevertheless, GFP and Cy5 signals remain highly correlated on a pixel-by-pixel basis after both rounds of fluorescence labeling ($r > 0.95$).

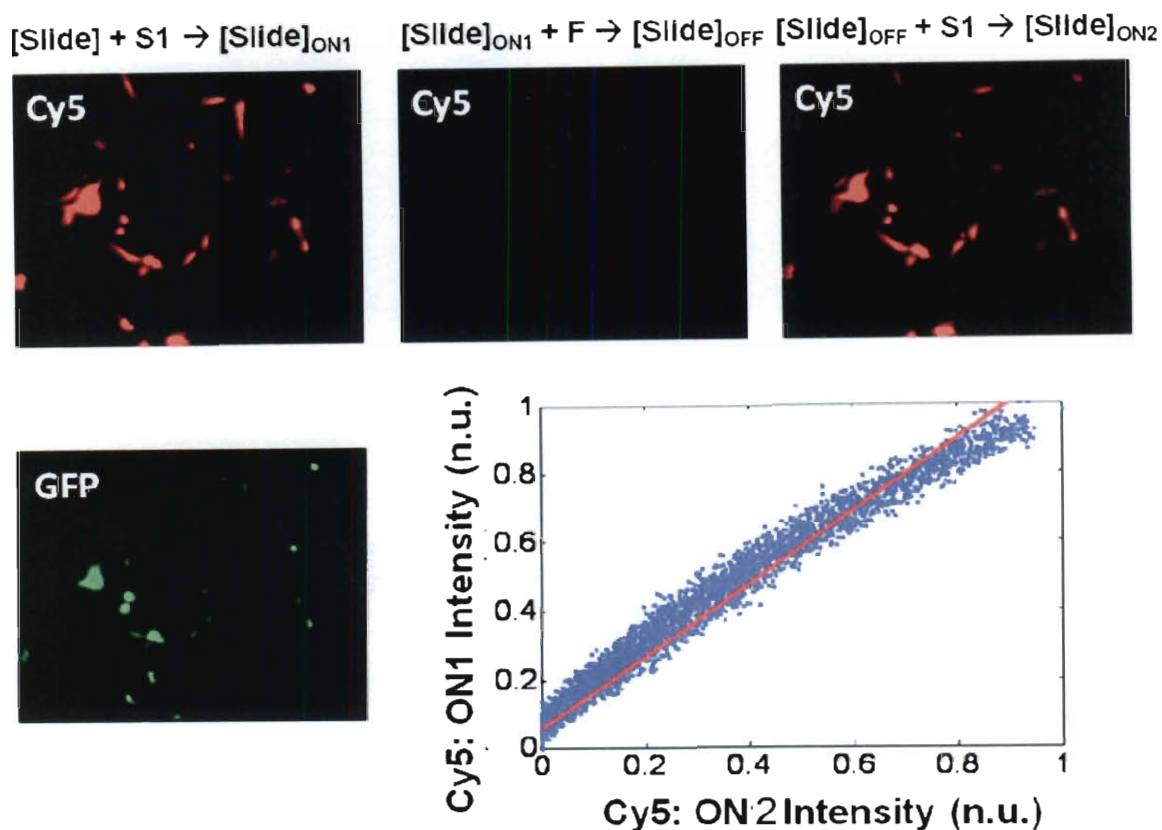


Figure 3.9: Reiterative circuit labeling of GFP-ZE transfected cells.

Sequential images of GFP-ZE transfected cells where circuit complexes were used to label (ON1), erase (OFF), and relabel (ON2) the same sample of HeLa cells are shown. The pixel intensities in both Cy5 images (Cy5/ON1 and Cy5/ON2) are linearly correlated; fitted slope = 1.1, $r = 0.92$.

Furthermore, strong correlations were found between cell images collected after the first and second dye labeling reactions are performed (i.e., between the ON1 and ON2 images in

Figure 3.9). The Cy5 intensities of both images are linearly correlated and can be approximated by a line possessing a slope of 1.1. We therefore conclude that the Cy5 dyes can be coupled to the GFP-Z_E markers with near-identical efficiencies through sequential rounds of circuit-based labeling.

3.4. Conclusion

In summary, we have demonstrated that DNA circuit complexes can be used as erasable molecular imaging probes. Here, the sequence-dependent specificity of the DNA circuit reactions facilitates multiplexed marker detection. The use of strand displacement mechanisms also allows fluorescence reporting complexes to be disassembled, and hence, new reporting complexes can be created and used to image additional sets of molecular markers. Importantly, these reactions can be carried out at ambient temperature and in mild buffering conditions to minimize potential perturbations to a biological specimen. While such capabilities should offer opportunities to increase the number of molecular markers that one can examine on a single biological sample via fluorescence microscopy by at least a factor of 2 or 3, the next challenge will also be to develop diverse sets of molecular targeting agents (e.g., monovalent DNA-conjugated antibodies) that can facilitate efficient molecular marker recognition and react with the DNA circuitry reliably. Efforts to optimize syntheses of such agents are currently underway.

Chapter 4

Optimized On-Cell DNA Strand Exchange Reactions for In Situ Marker Analyses³

This chapter provides an assessment and comparison of three different classes of imaging probes (probe constructs) to elucidate the contributions of complex size, free energy differences between states and strand displacement reactions on the labeling and erasing efficiencies on fixed cells. The kinetic rates of reaction of the three classes of imaging probes are compared to their calculated net free energy of reaction to determine which factor results in a more efficient and effective probe design. The chapter concludes with the development of an optimized dynamic nucleic acid probe system based on the results obtained from the characterization of the three probe constructs with short reaction times, non-prohibitive cost of production, highly efficient labeling and effective erasing steps.

³ The content of this chapter is being prepared for publication. The following authors have contributed Duose D. Y, Schweller R. M., Zimack J, Rogers A. R, Hittelman W. N and Diehl M. R.

4.1. Introduction

Recent advances in the field of DNA nanotechnology have facilitated the creation of various dynamic DNA and RNA complexes that can function effectively as programmable logic gates⁴¹⁻⁴³, chemical amplifiers⁴⁴⁻⁴⁵, and reconfigurable molecular structures⁴⁶. The unique potential of dynamic nucleic acid complexes for molecular-cell analyses are beginning to be realized. Engineered RNA hairpin devices have been used as ‘smart’ therapeutic technologies that can both detect genetic disease indicators *in vitro* and, in response, produce double stranded oligonucleotide polymers that selectively trigger cell apoptosis⁵³. Multiplexed (5-color) *in situ* detection of mRNA transcripts in fixed drosophila embryos has also been demonstrated⁵². Yet, despite such advances; the translation of dynamic oligonucleotide complexes towards these types of biological problems remains generically challenging. Most candidate probe constructions are first evaluated in a test-tube where displacement reactions occur in homogenously-mixed solutions⁶². The environment inside cells is much more complex and heterogeneous, and whether or not a sample is fixed and permeabilized, issues surrounding the sample penetration and probe dispersion must now be addressed. Other environmental factors may potentially interfere with the strand displacement process, and such effects could result in unwanted reverse or side reactions. Consequently, addressing the unique challenges of detecting bio-macromolecules within cells and tissues requires characterization of the strand exchange processes within these environments, and potentially, alterations to the designs of existing DNA systems to produce effective molecular probe systems.

Effective molecular probe system must meet the following design criteria; high specificity of targets to respective probe complexes; efficient labeling and erasing to ensure that fluorescent signals can be used to fully quantify target abundance without the interference of signals from previous rounds of labeling, and short reaction times to allow for multiple rounds of processing on the same sample without loss of integrity.

Herein, we evaluate the labeling / erasing efficiencies and on-cell kinetics of three dynamic DNA probe constructions. Construct 1, a 3-stranded probe complex utilizes a 3-way strand displacement mechanism for both labeling and erasing steps. Construct 2, a 2-stranded probe complex also utilizes a 3-way strand displacement mechanism for both labeling and erasing steps. The third and final construct, construct 3 is a 2-stranded probe complex that utilizes a 3-way strand displacement labeling reaction and a 4-way strand, displacement erasing reaction.

We show that efficient marker labeling and erasing is not necessarily dependent on the free energy differences between the different metastable states of the dynamic DNA complexes. Instead, non-toehold mediated strand displacement events can trigger reverse reactions (e.g., relabeling reactions) that reduce the overall rates that fluorophores can be removed from the sample. For probes that utilize a 3-way strand displacement mechanism, such control can be achieved via the cooperative binding of two distinct ssDNA components in order for a reverse reaction to occur.⁴⁸ Alternatively, fluorescent dyes can be removed efficiently using probe designs that exchange their nucleotides via 4-way strand displacement processes. As opposed to three-way systems, the erasing reactions with these complexes do produce inert (non-reactive) dsDNA products; the initiation rates of non-toehold mediated strand exchange are significantly lower. While

facilitating the optimization of probe designs for our application, such information should aid the development and implementation of other dynamic DNA systems that are designed for analogous future biological applications.

4.2. Materials and Methods

Oligonucleotides were purchased from Integrated DNA Technologies (IDT). The protein target, a recombinant autofluorescent protein, GFP-Z_E, was produced using standard cloning and cell transfection procedures. The C-terminal leucine zipper (Z_E) is used as an affinity tag for DNA labeling. This zipper associates strongly ($K_D \sim 10^{-15}$ M) with a complementary basic zipper (Z_R) that is incorporated into a DNA-conjugated artificial protein (Z_R-ELS₆-ssT) as a gene fusion⁶³. These polymers were produced according to the procedures presented in reference 56. The GFP-Z_E offers some advantages as a protein target for the present study since it can be outfitted stoichiometrically with a single ssDNA using conjugates that are purified to homogeneity; thus, reducing problems that could stem from batch to batch variability associated with DNA conjugation of conventional protein targeting agents.

4.2.1. DNA Probe Design

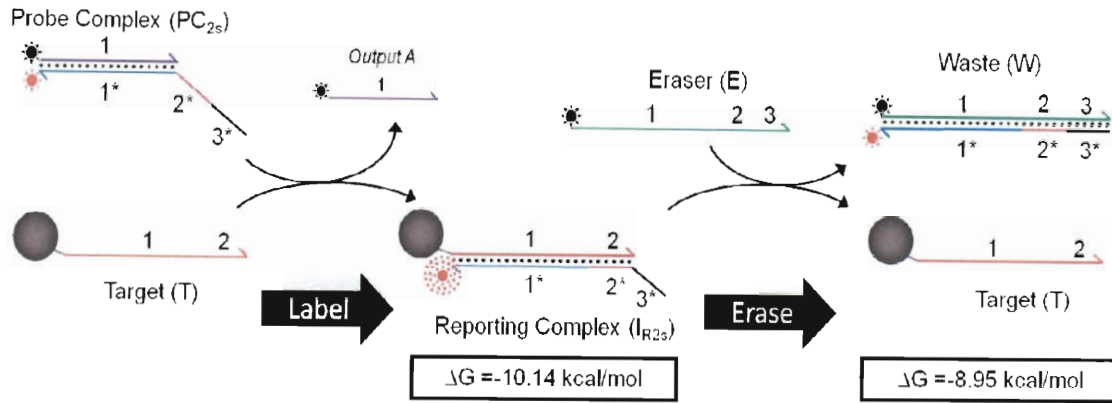


Figure 4.1: Labeling and Erasing Displacement Reactions using a 2-strand Probe complex.

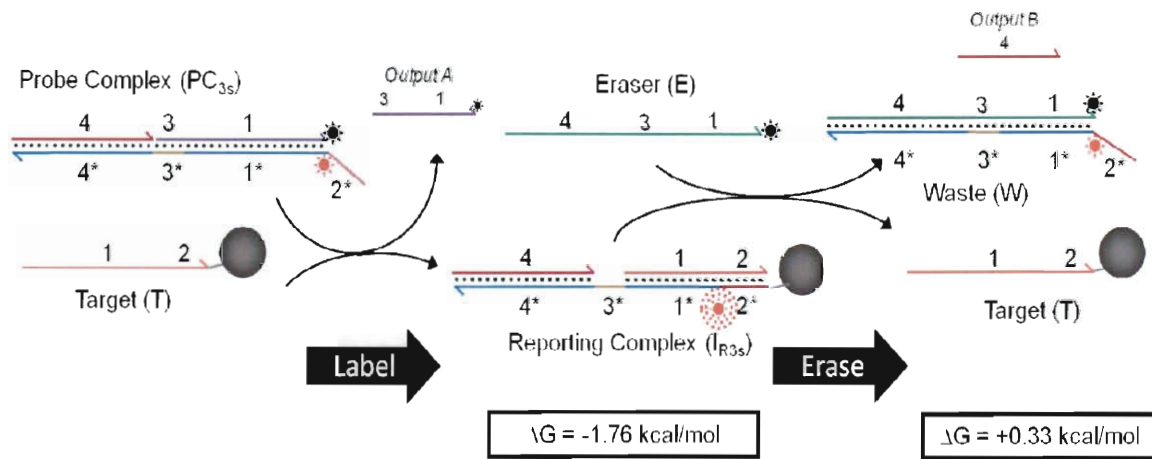


Figure 4.2: Labeling and Erasing Displacement Reactions using a 3-strand Probe Complex.

Probe sequences were designed using similar methods to those described in references 51 and 54. All sequences, excluding those taken directly from Zhang et al., were selected using a custom MATLAB script that generates random domains of specified lengths having pre-determined GC% range, while excluding previously generated domains or other prohibitive sequences (i.e. G quadruplexes), and avoiding secondary structures (i.e. hairpins). The generated domains (i.e. toehold domains, specificity domains and complementary domains) are ranked according to their normalized two-state hybridization energies with existing probe strands using algorithms from mFold³¹. The domains are then screened through the BLAST database⁶⁴ to minimize sequence homology with the mRNA transcriptome. The optimal domains are then selected manually from this list and concatenated with other domains to create full oligonucleotide sequences that will be incorporated into a probe complex. Other global criteria such as temperature, strand concentration, salt concentration, and the incorporation of GC clamps are specified prior to domain design. For simplicity, the contributions of dyes and quencher molecules were neglected in the free energy calculations.

Fluorophores (Cy3 or Cy5) and quencher molecules (Iowa Black) were incorporated in opposing strands at positions that minimized their intermolecular distance.

4.2.2. Cell Labeling and Erasing Procedures

CHO cells were grown on glass coverslips in DMEM supplemented with 10% FBS. After 24 hours, the culture medium was replaced and the cells were transiently transfected with vector containing the GFP-Z_E construct using Fugene (Roche) according to the manufacturer's protocol. Cells were cultured for an additional 12 hrs to allow for GFP-Z_E production. The cells were then fixed using freshly prepared 4% paraformaldehyde for 30 min. Activated aldehydes resulting from the fixation procedure were quenched using 1 mg/ml NaBH₄ for 5 min at room temperature. Afterwards, the cells were permeabilized using 0.2 % Triton X-100, washed twice with PBS, and stored overnight at 4 °C.

Prior to cell labeling experiments, the coverslips were rinsed twice in PBS, dried under an airstream, and then affixed to custom-fabricated micro-well chambers (10-round wells with 0.36 cm² culture area and culture volume of 400μl) using a precision-cut double sided adhesive film. The cells were re-hydrated with PBS before proceeding with labeling procedure. To minimize non-specific binding of the Z_R-ELS₆-ssT, cells were first blocked for 2 hrs using a solution containing 1 % BSA, 1 mg/ml Herring Sperm DNA and 0.5 uM polyT DNA in PBS. The cells were then incubated with a 400 nM solution of Z_R-ELS₆-ssT for 2 hours, and washed twice with PBS.

GFP-Z_E -labeling experiments were performed by incubating cells with solutions of 100 nM probe complexes in (TAE buffer supplemented with 12.5 mM Mg⁺²). Probe deactivation / erasing reactions were performed using 1 μM of the eraser strands (E_s) or complexes (E_c). All cell labeling / erasing reactions were carried out for 2 hours at 30 °C

using a rotating incubator shaker (200 rpm), except for the kinetic experiments where the reactions were performed directly on the microscope at room temperature and without shaking.

Cells were imaged using an inverted Nikon microscope outfitted with a 40x 0.9 N.A. objective, electronic shutters, and cooled EMCCD camera (Luca; Andor). A mechanical transition stage and electronic focusing mechanism was used to collect 5-10 different image fields for each sample. Images were processed using Nikon (NIS image) or ImageJ software³², and are presented as heat maps since this rendering enhances the contrast of low-level, remnant fluorescence signals within the ‘erased’ images.

4.3. Results and Discussion

4.3.1. Dynamic DNA Probe Designs

Our previous report showed that a class of 3-strand DNA complexes (PC_{3s}) developed originally by Zhang *et al*⁵⁴, can function effectively as erasable molecular imaging probes. Yet, despite this success, we believe our application would benefit from the development of somewhat smaller, 2-component probe complexes that incorporate terminal (3’ or 5’) dye molecules instead of the internal dyes in our prior designs; which restricts the types of dyes that can be incorporated into a probe⁶⁵. To address this issue, we created several DNA probe complexes composed of two partially complementary DNA oligonucleotides (PC_{2s} ; Figure 4.1). As with the 3-strand DNA complexes (Figure 4.2), these probes react with their ssDNA targets (T) via toehold-mediated strand exchange to produce a fluorescent reporting complex (I_{R2}) containing an unquenched fluorophore;

thus, molecular targets that are outfitted with their ssDNA targets can be visualized using fluorescent microscopy (Figures 4.3 and 4.4).

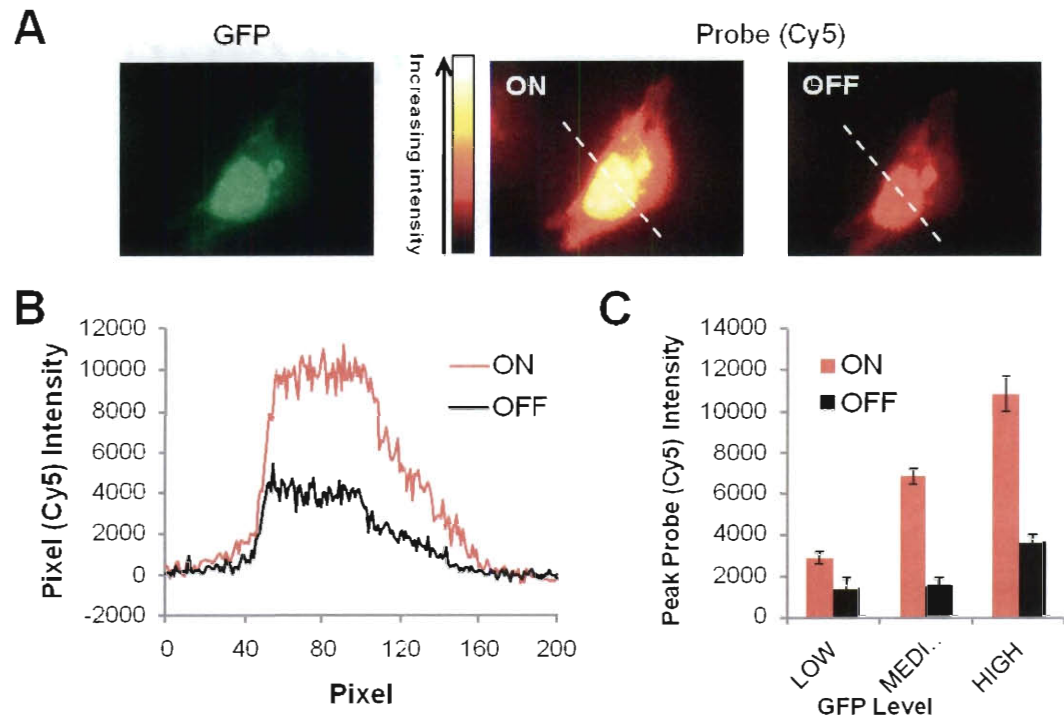


Figure 4.3: Labeling and removal of Cy5 fluorophores from protein markers via strand displacement reactions of a 2-strand probe complex.

(A) Selective labeling of expressed GFP proteins in CHO cells. Strong correspondence between the GFP and 2-strand probe (Cy5) signals is found; however, 20% of active Cy5 dye remains on the cells after the erasing reaction. (B) Pixel intensities for cross section indicated in the Probe images for both the ON and OFF states of the cells. (C) Histogram of the average Cy5 signal intensities for the ON and OFF states of 20 cells. Cells are grouped based on their GFP fluorescence intensities as described in the methods section.

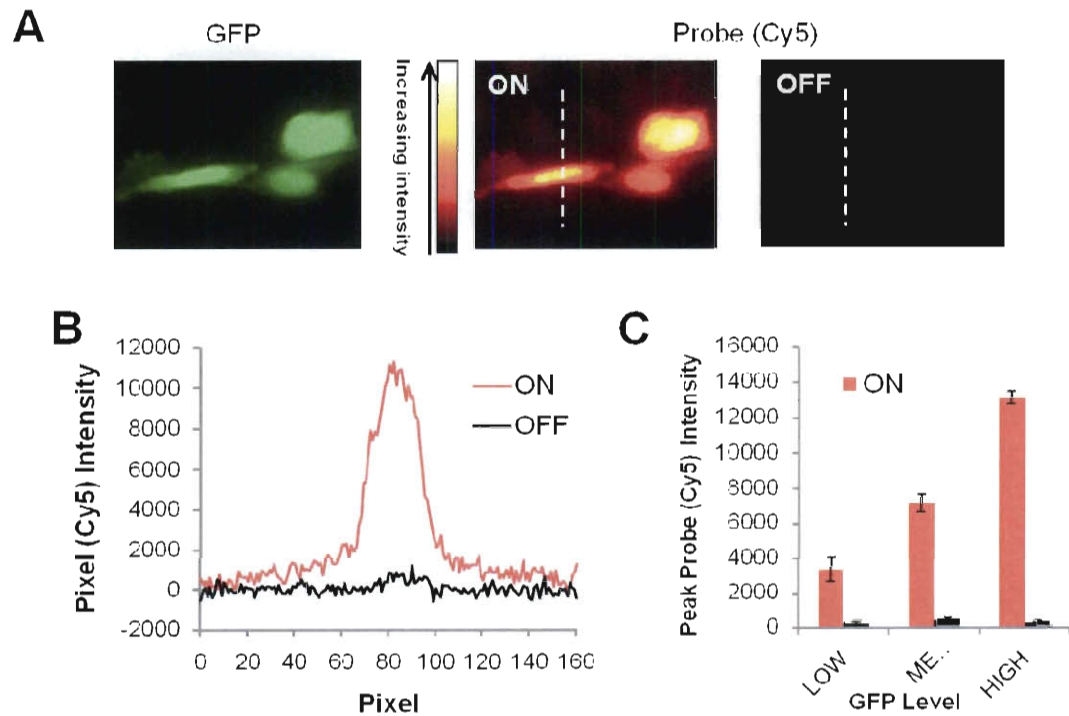


Figure 4.4: Labeling and removal of Cy5 fluorophores from protein markers via strand displacement reactions of a 3-strand probe complex.

(A) Selective labeling of expressed GFP proteins in CHO cells. The OFF reactions are now efficient, with S/B of 1.08. This result is reflected in (B) pixel intensities for cross sections as well as (C) histogram of average Cy5 signal intensities for 20 cells in their OFF states.

Analogously, the fluorescent reporting complexes for the 2- and 3-strand probe systems (I_{R2} and I_{R3} respectively) are disassembled using a single-stranded eraser oligonucleotide (E). Of note, the strand displacement process for the labeling and erasing reactions of each probe system proceed via a 3-way branched migration⁶⁶, where a ssDNA component displaces with one oligonucleotide in a duplexed complex.

The PC_{2s} probes contains three distinct domains: an 18 bp domain that is completely hybridized (domain 1), and two 6 bp toehold domains that are positioned adjacent to one another at one end of the complex (domains 2* and 3*). The reporting complexes of the PC_{2s} system are formed through a toehold-mediated exchange reaction that is initiated by hybridization of the probe complex to domain 2 of the 'target' strand. However, unlike the 3-strand complexes, whose reaction is completed by the release of an output strand from a second toehold domain (Output A and domain 3 in Figure 4.2), T displaces the output strand within PC_{2s} completely. This reaction therefore results in a net gain of 6 matched bases within I_{R2} ($\Delta G = -10.14$ kcal/mol). Similarly, the I_{R2} reporters are disassembled /erased via a reaction of I_R with an eraser strand (E) that binds to toehold 3* on I_{R2} and displaces the T from the reporting complex. This reaction produces a completely duplexed waste complex (W), and yields another gain of 6 matched bases ($\Delta G = - 8.95$ kcal/mol). In contrast, I_{R3} of the 3-strand system is disassembled / erased via a reaction where E binds to a 4 bp toehold, and displaces two strands (output B and T). In this case, T is released from the complex after it disassociates from a 6 bp toehold, and hence, the composite reaction of E with I_{R3} is energetically unfavorable ($\Delta G = + 0.33$ kcal/mol).

The net free energies were calculated using Nupack⁶⁷ software. The free energy of formation of the 3-strand complex was calculated using the DNA sequences that formed the complex and setting the temperature, sodium and magnesium ion concentrations to 25°C, 0.05M and 0.015M respectively in Nupack. The same criteria were used in the calculation of the free energies of the reporting complex, target strand, eraser strand and waste complex. The net free energy for labeling was calculated by subtracting the free energy of formation of the reporting complex from that of the substrate complex. The free energies of the catalyst and output strand were designed and calculated to be 0 kcal / mol.

4.3.2. Selective *In Situ* Labeling and Erasing of DNA-conjugated Proteins

As indicated above, both the labeling and erasing reactions of the 2-strand, PC_{2s}, probes system are more thermodynamically favorable than those of the 3-strand PC_{3s} probes. Thus, one may expect the PC_{2s} probes to function as more effective, erasable molecular imaging probes. However, this behavior is not found in cell imaging experiments. Here, exogenously expressed GFP proteins within fixed and permeabilized CHO cells were first outfitted with ‘target’ strands using an ssDNA-artificial protein conjugate and then reacted using dynamic DNA probe complexes that incorporate Cy5 fluorophores (Figures 4.1 and 4.2). In each case, labeling intensities were evaluated after the probes were allowed to react for a period of 2 hours; reactions will likely need to be completed within this time frame to implement our reiterative marker imaging technique.

Comparisons of GFP and Cy5 signals produced after a labeling reaction show that both the 2-strand and 3-strand probe constructs can selectively couple fluorophores to DNA-conjugated proteins on cells. In each case, the untransfected cells are not labeled and the microstructural organization and the intracellular distribution of GFP molecules are reproduced in the Cy5 image within the transfected cells. Near indistinguishable pixel-by-pixel correlations are found between GFP and Cy5 signal intensities with both types of probe constructions. Yet, despite their similar performance with respect to protein labeling, the erasing reaction of the PC_{2s} probe system is found to be much less efficient than the PC_{3s} system (Figure 4.3). Here, the erasing reaction that disassembles the 2-strand I_{R2s} reporter results in residual (Cy5) signals that are approximately 20% of the signal amplitudes produced by the prior marker labeling reaction. Furthermore, these unwanted 'OFF-state' signals are positively correlated with both the GFP levels and the probe-generated 'ON-state' Cy5 intensities (Figure 4.3C). This behavior is not found for the 3-strand complexes (Figure 4.4). As with our prior work, 'OFF-state' signals with the 3-strand complexes can barely be detected over background autofluorescence of the slide (S/B =1.08). Thus, of the two different probe systems evaluated thus far, only the 3-strand systems can function effectively as an erasable molecular imaging probe.

4.3.3. On-Cell Kinetics of Strand Displacement Reactions

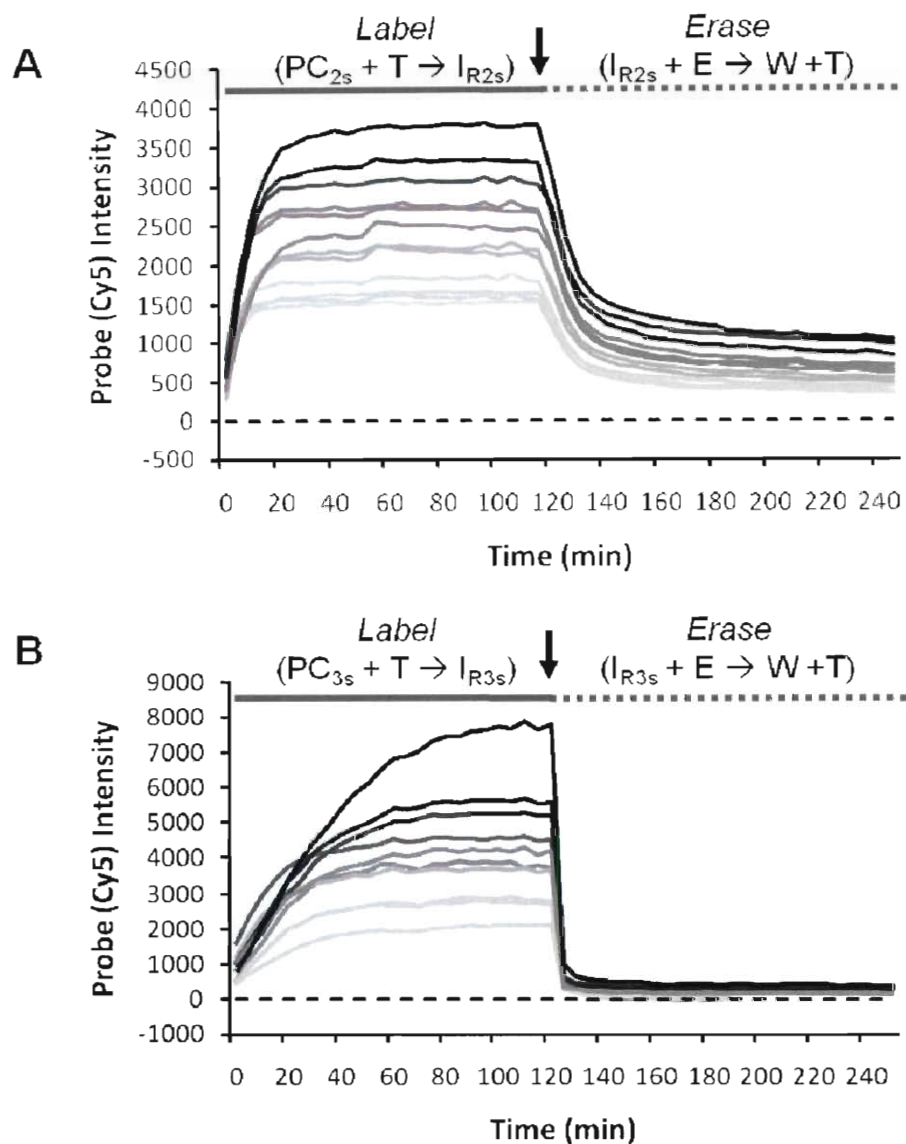


Figure 4.5: Kinetics of DNA strand Displacement reactions on fixed cells

(A) Labeling and erasing reactions of a 2-strand complex. The erasing reactions are inefficient and do not go down to background levels. (B) Labeling and erasing reactions of a 3-strand probe complex. Both the labeling and erasing reactions are efficient with the erasing reaction going down to background signal levels.

We next characterized the on-cell marker labeling and erasing kinetics of the PC_{2s} and PC_{3s} systems by monitoring the rates that probe (Cy5) signals co-localized to, and then could be removed from, expressed GFP molecules (Figure 3.5). For these experiments, cells were imaged every 5 minutes during the marker labeling / erasing procedures. These experiments show that GFP markers are labeled rapidly by the 2-strand probes, and that probe signals saturate within 20 min, even for the cells possessing the highest GFP expression levels. Although the PC_{3s} labeling reactions were somewhat slower, Cy5 intensities reached their saturated values in less than 1 hour, except for the very brightest cells within the sample. Thus, both probe complexes can support efficient (relatively fast and proportionally quantitative) marker labeling. Yet, there are significant differences in the erasing kinetics of these systems. Although positive Cy5 fluorescence intensities decrease rapidly after the initial addition of E in both cases, the PC_{2s} system erases more slowly than the PC_{3s} system. In addition, and, more importantly, the PC_{2s} erasing reaction slows appreciably after a period of ~20 minutes, and significant Cy5 signals (20%) remain of the sample after the full 2-hour incubation period. In contrast, signals in the 3-strand complexes drop rapidly to less than 3% of background signal levels.

As mentioned above, given the larger free energy difference between the PC_{2s} and its corresponding I_{R2s} reporter, the labeling reaction of the 2-strand probe system is more energetically favorable than that of the 3-strand system. Thus, observations of higher labeling / 'ON' rates when PC_{2s} is reacted with T are consistent with the differences in the thermodynamic properties of the different probe constructions. However, the fact that the erasing reaction for the 2-strand probes system is also thermodynamically favorable,

while the corresponding erasing reaction of the 3-strand probes is unfavorable, suggests that some other factor must be influencing the strand-exchange process during the erasing process. One possibility is that the reaction between E and 3-strand reporter I_{R3s} produces a metastable intermediate complex where the fluorophore-containing strand is still coupled to T, but rendered inactive by the quencher. This complex could exist if T does not fully release from W during the erasing reaction (Figure 3.6).

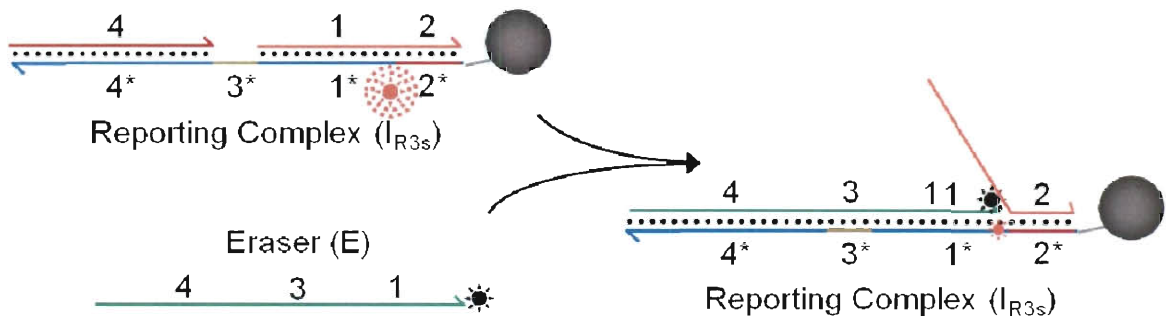


Figure 4.6: Intermediate Complex Formation during Erasing Displacement Reactions with a 3-strand probe complex.

In addition, such a reaction could be thermodynamically favorable since there would be a gain of 4 matched bases due to the hybridization of E (domain 3) to the 3* toehold in I_{R3s} . To test this, the on-cell kinetics for both the 2- and 3-strand probe complexes were evaluated using eraser strands that do not contain quencher molecules (Figure 4.7); thus, duplexed I_R , other unidentified intermediate, and W complexes will still produce fluorescence signals if they remain bound or trapped within the cells. In this case, the

kinetic curves measured for unquenched 3-strand erasing reaction: $I_{R_{3s}} + E \rightarrow W + \text{output}$ 2 are very similar to the original plots in figure 4.5B. Thus, these results suggest W is able to release from T in this reaction, and hence, the ability for the PC_{3s} system to erase efficiently is not derived from the formation of the intermediate state complex depicted in Figure 4.6. Furthermore, the rapid drop of Cy5 signals also indicates that the (now fluorescently active) W complex is 'inert' (non-reactive and unbound to T) and can freely diffuse out of fixed cells.

Similarly to quencher-bearing E, the reaction of 2-strand reporter complexes $I_{R_{2s}}$ with non-quencher bearing E leaves appreciable remnant Cy5 signals on GFP transfected cells due to incomplete erasing (Figure 4.7 top). However, in this case, very intense Cy5 signals remain on the sample; 'OFF-state' pixel intensities are approximately 75 % of their 'ON-state' values. Of note, T does not release from the waste complex through toehold de-hybridization in this reaction (T is displaced completely from the complex), and consequently, analogous intermediate state complexes to the metastable intermediate depicted in Figure 3.6 are not expected to influence the PC_{2s} system's erasing kinetics.

While the kinetic analyses of this erasing reaction without quenchers indicate significant numbers of $I_{R_{2s}}$ reporters remain on the cells after 2 hours, this result also suggests that fluorescently active W complexes are somehow bound to target and trapped within the cells' volume. Such behavior is somewhat surprising given the high concentration of E (1 μM) and reaction volume (100 μL) used for the erasing experiments. With such an excess of E relative to the total number of $I_{R_{2s}}$ complexes on the cells, there should be a strong driving force to push the erasing reaction forward.

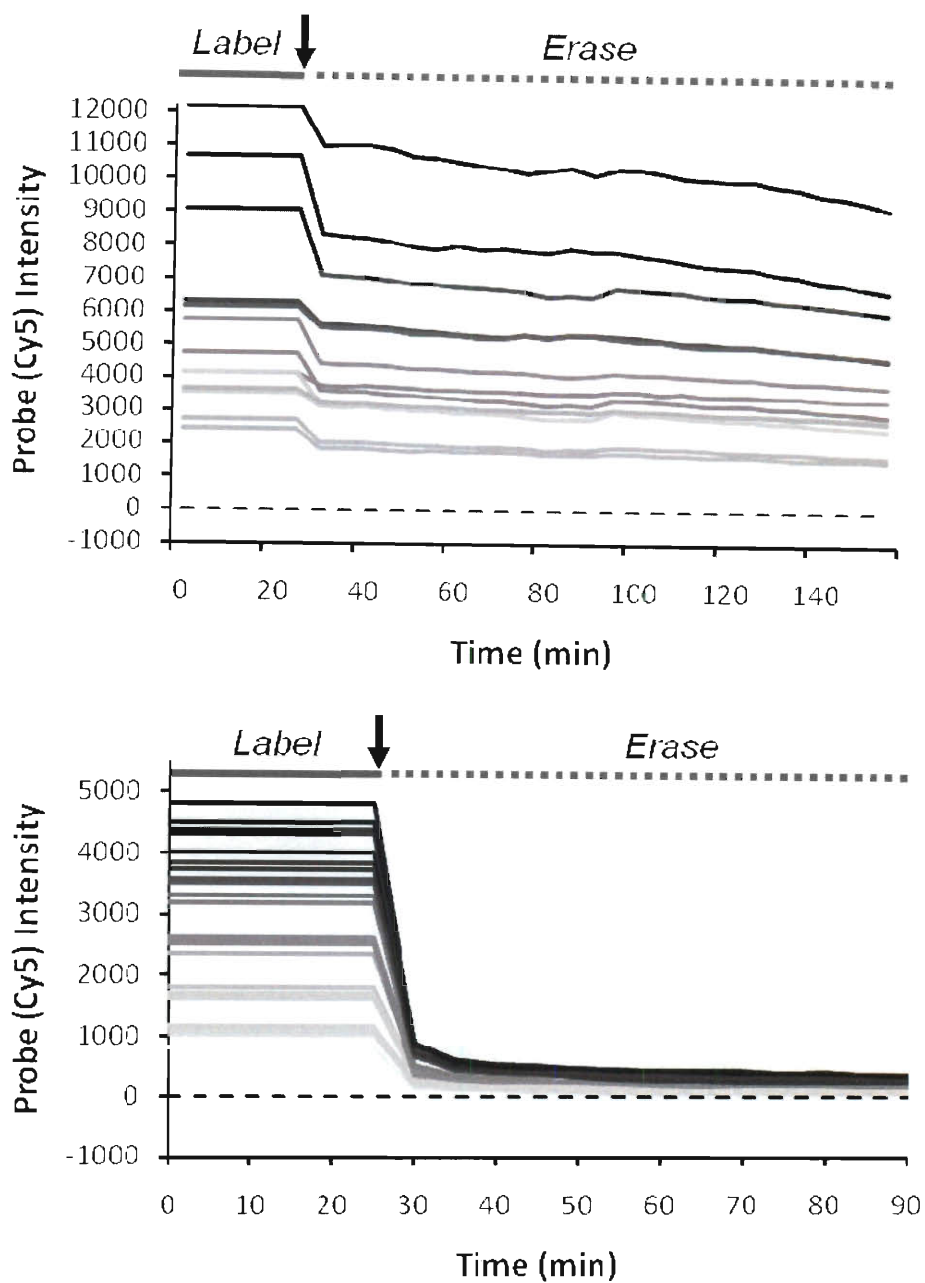


Figure 4.7: Labeling and Erasing Kinetics of non-quencher bearing Eraser strand

(Top) 2-strand complex showing about 75% residual signal (Bottom) 3-strand complex with low residual signal.

However, given the compartmentalization of these targets within the cells, the local concentration of T and W can be quite high, which could serve to drive the reverse (relabeling) reaction: $W + T \rightarrow I_{R2s} + E$. Such effects could, in turn, influence the rates that the fluorophores within the W complexes are released from the cells since the strands containing the fluorophores will now transiently interact with multiple immobilized T strands as they diffuse towards the cell boundary.

Interestingly, W of the PC_{2s} system does not contain a toehold, and thus, the reverse / relabeling reaction must occur via a non-toehold mediated exchange process. Of note, these types of reactions have been shown to be enhanced when molecular crowding agents are present in solution⁶⁸. Consequently, it is possible that proteins and other macromolecules within the cells act similarly and potentially also accelerate the rates of non-toehold mediated strand exchange.

The observation that the W complex of the 3-strand probe system is released efficiently from the cells further supports the notion that reverse (relabeling) reactions reduce the erasing efficiencies of the 2-strand systems. In this case, two separate ssDNA must bind to the W complex simultaneously in order to produce the fluorescently active I_{R3s} reporters. Thus, W complexes are less likely to re-associate with T and can more readily diffuse out of the cells. For our application, such control appears to be necessary to produce probes that erase efficiently while using the 3-way strand displacement mechanism.

4.3.4. Efficient Signal Erasing via Four-way Branched Migration Reactions

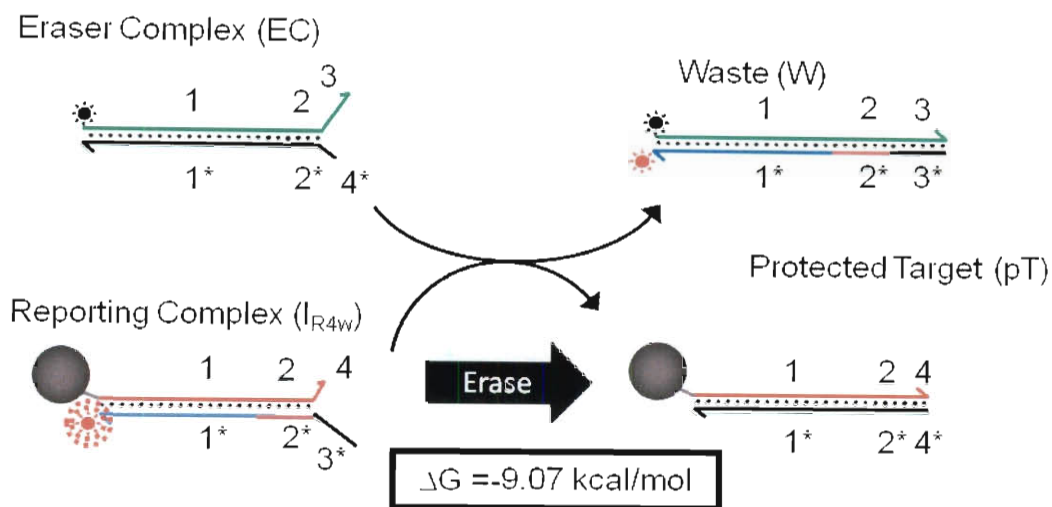


Figure 4.8: Erasing Displacement Reactions using 4-way branch migration in a 2-strand probe complex.

Based on the notion that non-toehold mediated exchange of oligonucleotides within W complexes reduces the erasing performance of the PC_{2s} probes systems, we designed a simple 2-strand probe complex. This complex, will allow the use of terminal (5' or 3') dye molecules and quenchers, and will produce W products that are more 'inert' and can diffuse out of the cells more rapidly (Figure 4.8). To do so, we modified our original 2-strand complexes such that they bind the same sequence T, but leave a second 2 bp toehold un-hybridized within the I_R complex (domain 4 in Figure 4.8). In this case, the strand displacement process driving marker labeling is nearly identical to that of the original PC_{2s} systems. However, the erasing reaction now proceeds via a four-way branched migration and exchange process⁶⁹, where two different toeholds assist in the

initiation of the reaction. Since this reaction now produces a fully duplexed W complex, we anticipated that this modification would overcome the problems associated with the reverse (relabeling) reaction that prevents complete marker erasing in the PC_{2s} system. While this system produces near identical labeling performance as the original PC_{2s} probes (Figure 4.2B), and is at least as fast as the original system, the erasing step of the procedure is also efficient with this modification, yielding 'OFF-state' intensities and reaction rates that are near identical to those of the PC_{3s} system (Figure 4.9). Consequently, this probe construction appears to possess all of the attributes required for our reiterative marker detection procedure.

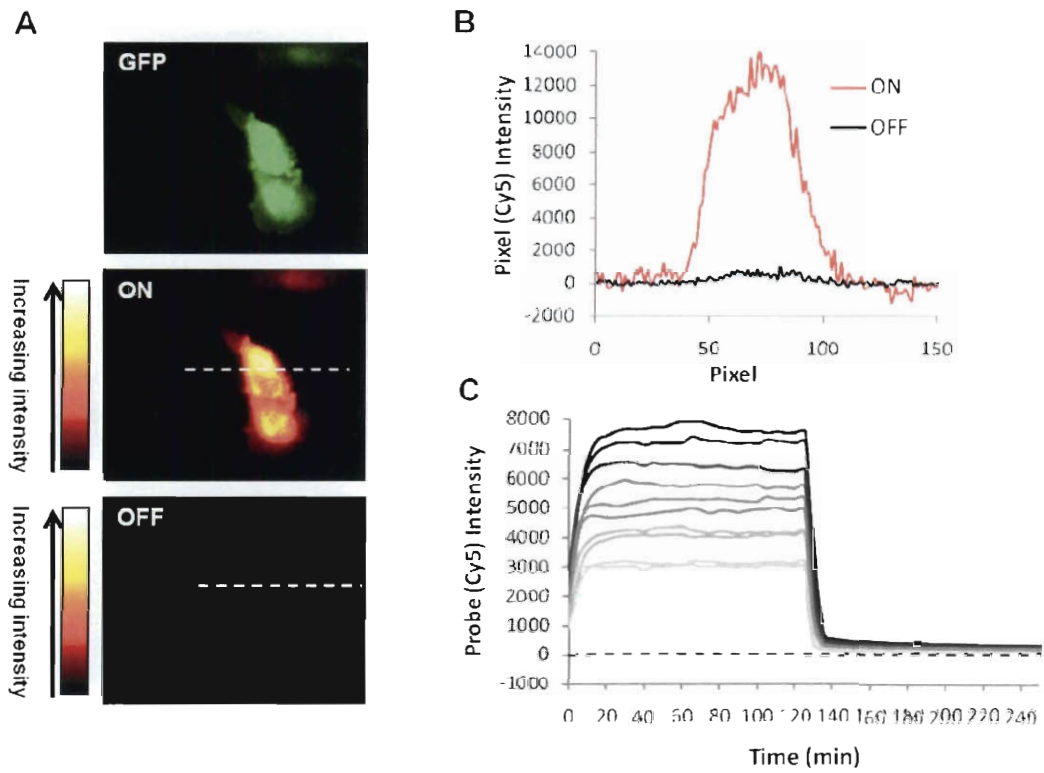


Figure 4.9: Labeling and removal of Cy5 fluorophores from protein markers using the 4-way-2-strand probe complex.

(A) Selective labeling of expressed GFP proteins in CHO cells. (B) pixel intensities for cross sections as well as (C) Kinetics of DNA strand displacement reactions on fixed cells showing the efficient labeling and erasing with 4-way-2-strand probe complex.

Chapter 5

Conclusion

This chapter provides a summary of the development and characterization of dynamic nucleic acid complexes as imaging probes for erasable multicolor imaging of molecular biomarkers in cells. The implications of the results obtained from this work and its applicability is discussed. The chapter concludes by providing a brief summary of results from on-going experiments that give new insights into the applicability of MRMBI to expanding the utility of immunofluorescence imaging.

5.1. Summary and Conclusions

The diagnosis and staging of cancer and other diseases is becoming increasingly reliant on the quantitative analysis of phenotypic indicators, called 'biomarkers', in histological samples. However, immunohistological methods remain substantially

restricted by the fact that only a few biomarkers can be examined on a single biological sample. The use of more than one specimen can circumvent this problem in some cases, but whenever heterogeneity across the specimens from a tissue confounds its characterization or is of central importance, diagnostic information is lost. The number of markers that can be evaluated on an individual biological sample could be increased if it were possible to remove fluorescent probes from cells such that new markers could be labeled and detected using the same fluorescent reporting molecules. However, fluorophores are typically attached covalently to probes that are engineered specifically to bind their targets with high affinity. As such, probe removal generally involves the use of harsh chemicals and/or physical treatments that can disrupt cell and tissue morphology and compromise subsequent marker analyses.

The goal of this project was to surmount this problem by developing and characterizing an erasable, multi-color imaging technology capable of detecting large numbers of molecular markers on a single biological sample. The technology utilized dynamic nucleic acid complexes as probes for the selective coupling and removal of fluorophores to and from biomarkers via strand displacement mechanism (i.e. the selective exchange of single oligonucleotide strands between complexes of DNA). The system consisted of (1) 'targets', which are ssDNA or partially hybridized oligonucleotide strands conjugated to a protein of interest for biomarker recognition in cells, and (2) multi-strand, fluorophore-containing DNA 'probe complexes' that react with the DNA portion of the target in a sequence dependent fashion via a toehold mediated strand displacement reaction to create fluorescent reporting complexes. The removal of the dye (erasing step) followed via the addition of a quencher-bearing single

or partially hybridized DNA complex that displaced the target's DNA strand in a sequence specific manner to ensure efficient removal of signal without perturbing sample integrity. By combining dynamic DNA complexes that operate independent of one another with spectrally-separated dyes, multiple and unique target/probe pairs that associate specifically were created imaged in parallel.

In order to successfully employ this technology for multiplexed and reiterative molecular biomarker imaging, the following design criteria must be met; high specificity of targets to respective probe complexes; efficient labeling and erasing to ensure that fluorescent signals can be used to fully quantify target abundance without the interference of signals from previous rounds of labeling, and short reaction times to allow for multiple rounds of processing on the same sample without loss of integrity. To meet these goals three classes of probes were designed and their structure-function relationships elucidated to determine the contributions of complex size, free energy differences between states, and strand displacement (toehold and non-toehold as well as 3- and 4-way) on labeling and erasing kinetics and efficiencies on cells.

5.1.1. Summary of DNA Constructs Designed

Construct 1 consisted of a ssDNA target and a partially hybridized 3-stranded probe complex adapted directly from Zhang *et al*⁵⁴. This construct participated in a 3-way strand displacement reaction for both labeling and erasing steps. The calculated net free energy change was negative (-1.76 kcal/mol) for the labeling reaction and slightly positive (+0.33 kcal/mol) for the erasing reaction. Thus, the erasing reaction was expected to be slow and inefficient based solely on the net free energies. However, the

labeling and erasing kinetics on cells were fast and efficient. The signal to background ratio of the labeling step was 20:1 whilst that of the erasing step was 1.08:1.

Construct 2 consisted of a ssDNA target and a partially hybridized 2-stranded probe complex designed in-house that also participated in a 3-way strand displacement reaction for both labeling and erasing steps. This construct was designed to be small (about half the size of construct 1), with a large negative calculated net free energy for both the labeling (-10.14 kcal/mol) and erasing reactions (-8.95 kcal/mol). Therefore, based on the smaller size and net free energies, this construct was expected to have a faster and more efficient labeling and erasing reaction rate than the previous construct. Surprisingly however, although the labeling kinetics on cells was fast (reaching a saturation level earlier than in the previous construct), the erasing kinetics was slow and inefficient leaving behind a substantial amount of signal after a 2hr reaction time (same as for the previous construct) with signal levels approximately 20% of the labeling signal. The inefficient removal of signal in the erasing step was attributed to non-toehold mediated reverse reaction that occurred because the target strand was available to participate in reverse reaction.

Construct 3 consisted of a ssDNA target and a partially hybridized 2-stranded probe complex designed in-house that participated in a 3-way strand displacement labeling reaction and a 4-way strand, displacement erasing reaction resulting in the sequestration of the target strand making it unavailable to participate in the reverse reaction. This construct was designed to be small (similar in size to construct 2) to expedite diffusion in and out of the cell and to have a large negative net free energy for both the labeling (-9.98 kcal/mol) and erasing reactions (-9.07 kcal/mol). The expectation

for this construct was that the 4-way toehold mediated strand displacement reactions and sequestration of the target should speed up the erasing reaction on cells and result in a faster more efficient erasing step. As anticipated, the labeling and erasing kinetics on cells were fast and efficient with the efficiency of the erasing step similar to that of construct 1.

Thus, of the three constructs characterized, construct 1 and 3 met the metrics of the design criteria best with construct 1 having the added advantage of an available target strand that can participate in multiple labeling reactions of the same molecular marker with minimal signal accumulation.

This characterization showed that reaction efficiencies depend less on net free energy change than on the engineered state of the complex during the strand displacement reaction. Thus, the above characterizations provided a new paradigm for the optimization of dynamic nucleic acid complexes for fast and efficient erasable molecular biomarker imaging on a single biological sample.

5.2. Future Directions

5.2.1. Final probe design

A new and final design was proposed based on the information garnered from the structure-function relationship of the three probe constructs. An additional criterion of non-prohibitive cost of production was added to the design specifications previously mentioned. The new probe construct consists of a partially hybridized 3-strand complex

(with the third strand being a spectator strand that houses the quencher and serves to conserve quencher from construct to construct.) that participates in a 4-way strand displacement labeling and erasing reaction (Figure 5.1).

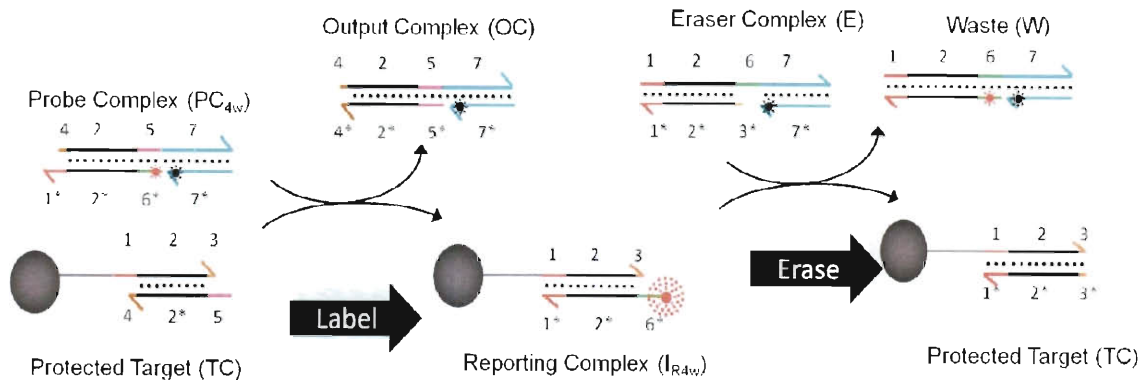


Figure 5.1: Labeling and erasing reactions using a 4 way-3-strand probe

This structure utilizes a conserved quencher region as part of the 3-strand complex and a partially hybridized duplex DNA in the DNA-protein conjugate so as to increase the labeling reaction by participating in a 4-way strand displacement reaction in the labeling step. It also uses a duplexed eraser complex similar to construct 3 that reacts in a 4-way strand displacement reaction to remove the dye from the target while sequestering the target strand thus preventing it from participating in any reverse reactions or subsequent labeling reactions.

Because of the conservation of quenchers in this construct, one quencher strand is used for both the erasing complex and the probe complex and this quencher strand is conserved among several different constructs with different targets thereby reducing the

cost associated with different quencher bearing strands for different targets. Consequently, barring any unforeseen complications, this design should function in a manner similar to that of construct 3 and all subsequent probe complexes will be designed according to the criteria set forth by this final construct design. Experiments are currently underway to determine the labeling and erasing kinetics of this final construct on cells.

5.2.2. Multiplexed imaging with an antibody binding protein: LG⁴

With the design and characterization of the final probe construct completed, the next phase of this project deals with the development of an engineered universal antibody binding protein that is small and has high affinity ($K_d=0.24\mu\text{M}$) to all antibodies⁷⁰. This universal antibody binding protein (protein LG), is fused to the negatively charged half of an engineered heterodimeric leucine zipper complex⁵⁶ to create a protein LG-zipper construct (LG-Z_E). This construct can then interact with the target portion of the dynamic DNA complex through association with the positively charged leucine zipper conjugated to the DNA (Z_R-ELS₆-DNA). The resulting protein LG-zipper-zipper-DNA ($\{\text{LG-Z}_E\} - \{\text{Z}_R\text{-ELS}_6\text{-DNA}\}$) construct can then be bound to any antibody of interest while maintaining probe complex specificity through the target's DNA sequence. To detect several markers on the same biological sample, individual antibodies are pre-bound to a

⁴ The data and information provided here for multiplexed imaging with antibody binding protein: LG is the work of Arthur Rogers PhD

specific protein LG-zipper-zipper-DNA complex in solution to form antibody-protein LG-zipper-zipper-DNA complex. This is to ensure that antibody-DNA complexes are formed stoichiometrically before being incubated on cells, so as to reduce crosstalk between antibodies and protein LG. Through this process, several different targets (ssDNA) can be self assembled with antibody-protein LG-zipper constructs as shown in Figure 5.2.

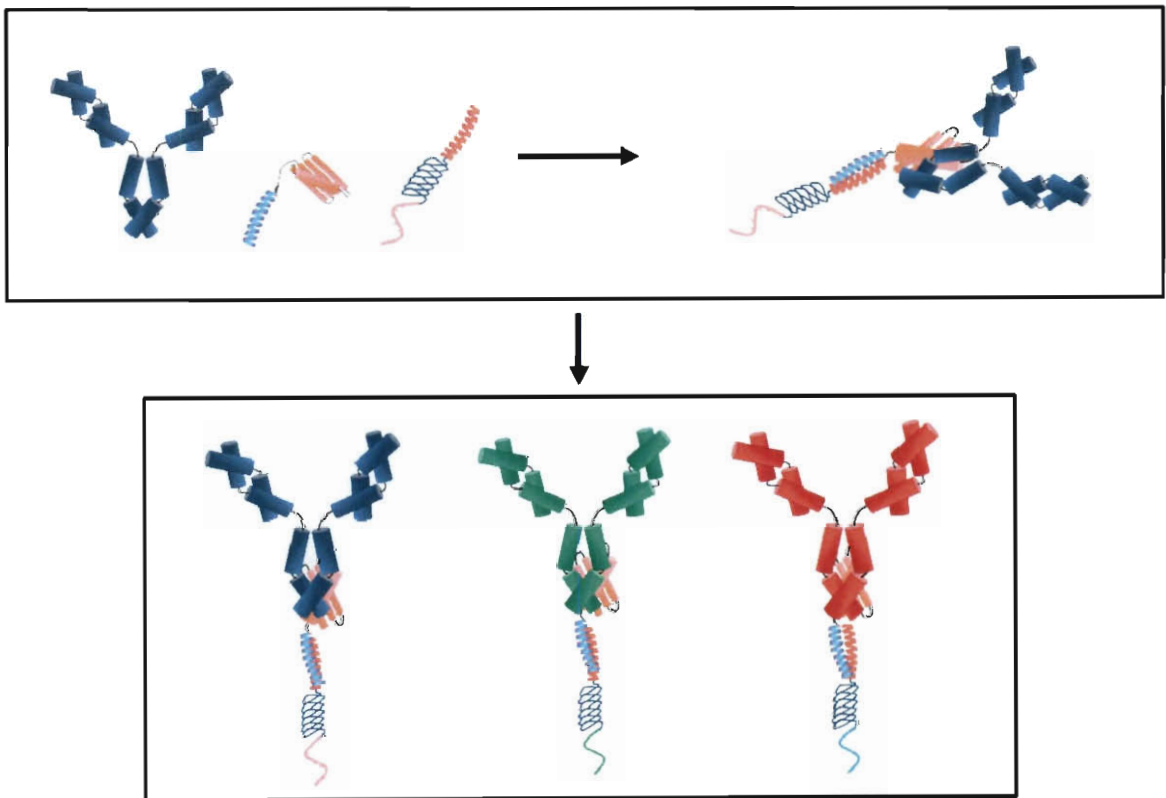


Figure 5.2: Scheme showing the self-assembly of leucine zipper -protein LG with ZR-ELS6-DNA and antibodies.

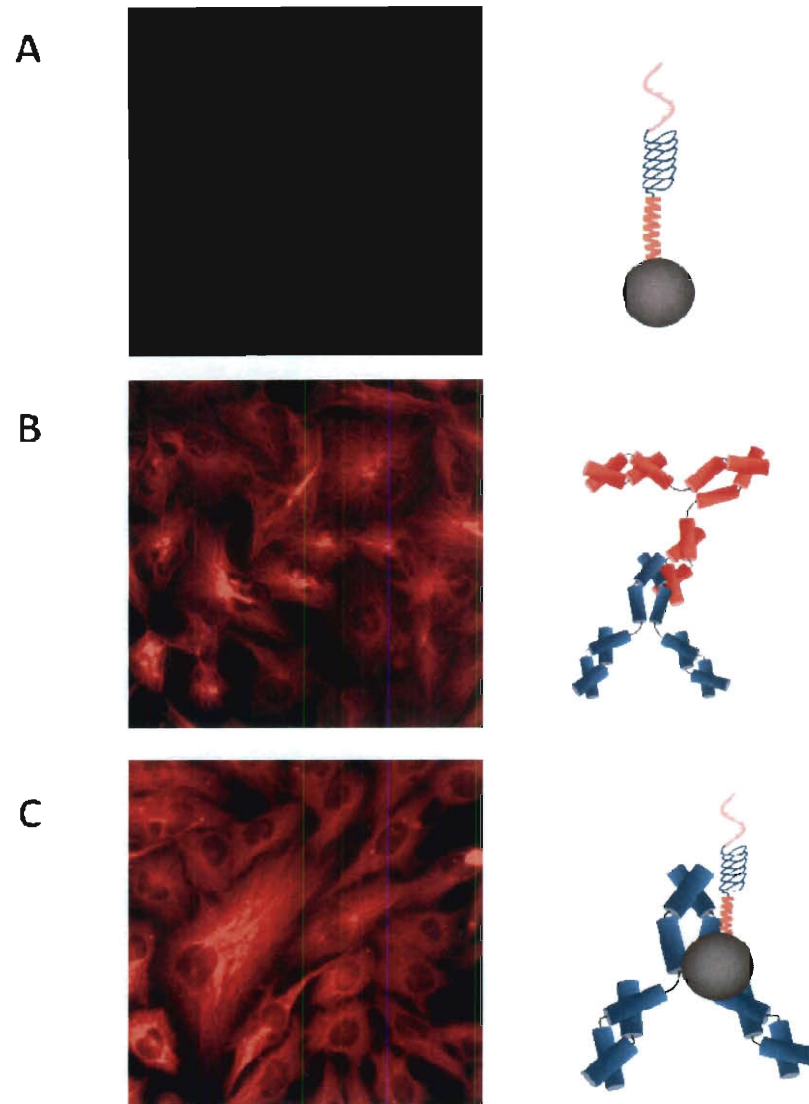


Figure 5.3: Comparison of immunofluorescence image of labeled MT

(A) protein LG-DNA complex without microtubule-antibody (B) primary microtubule-antibody and dye labeled secondary (C) protein LG-DNA complex with primary microtubule-antibody. All images were taken at 500ms exposure and contrasted the same.

Preliminary data from this endeavor is illustrated in Figure 5.3, where the labeling efficiency of microtubule with antibody associated to the protein LG-zipper-zipper DNA complex is shown to be qualitatively comparable to that of an indirect immunofluorescence labeling of the microtubules with primary and secondary antibodies. It is worthy to note that a control experiment with only the protein LG-zipper-zipper DNA complex (no antibody was associated to this complex) produced no signal implying low binding of protein LG to cells. Therefore this information provides preliminary evidence that a universal antibody binding protein (LG) can be used to image biomarkers in cells. Further experiments are still needed to ensure that minimum crosstalk occurs when multiple antibodies are used in a single labeling reaction.

5.2.3. Amplification

With the establishment of the self assembled antibody constructs in place, markers with low signal levels could be detected using dynamic DNA probe complexes with antigen recognition capabilities coupled with signal amplification. This signal amplification can be achieved based on the same strand displacement principle. Various dynamic DNA complex technologies can be integrated with the current probe design to achieve signal amplification. One such technology, Hybridization Chain Reaction (HCR)⁴⁴, can be integrated into the 3-strand probe construct to achieve signal amplification.

The HCR technology utilizes small DNA hairpins as substrates in a linear polymerization reaction (Fig. 5.4). Polymerization is triggered by the presence of a

specific ssDNA strand called the initiator. This reaction then produces a long linear polymer possessing multiple nicked duplexes.

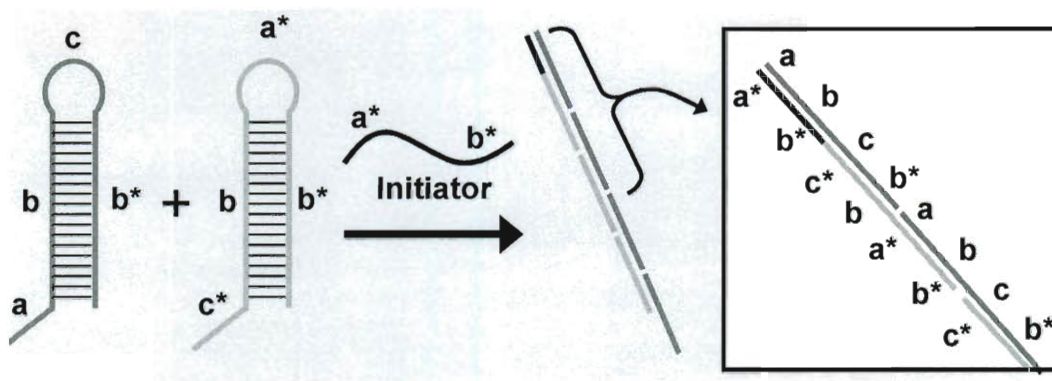


Figure 5.4: Hybridization chain reaction mechanism

The addition of the initiator to both hairpins triggers the formation of higher order polymers (the amplified product).

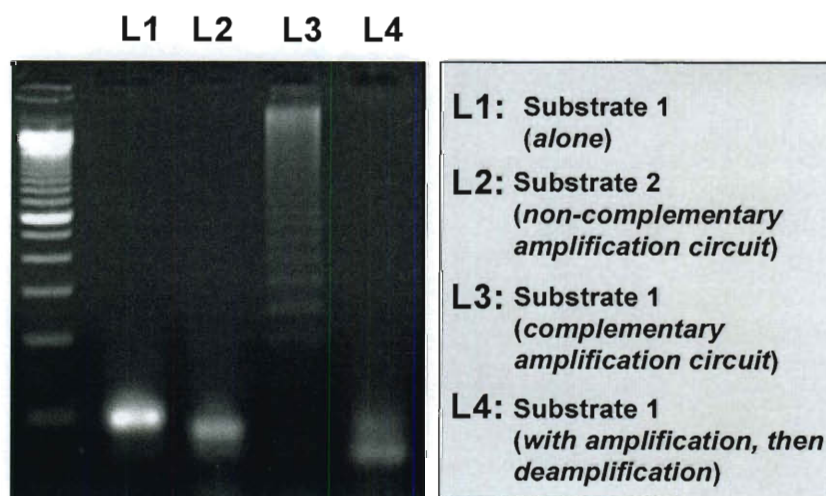


Figure 5.5: Native PAGE-gel showing the selective amplification and de-amplification (L4) of 3-strand probe construct via the HCR method.

L3 shows higher molecular weight products formed as a result of amplification.

Preliminary solution phase data in figure 5.5 shows that the addition of both hairpin strands to an initiator bearing 3-strand construct (substrate 1) resulted in the formation of higher order polymers (amplification) on the 3-strand construct. The addition of both hairpins to a non-initiator bearing 3strand construct (substrate 2) did not result in the formation of higher order polymers. This preliminary data shows that amplification schemes can be incorporated into MRMBI probes for detection of low level signals on cells and tissues. Further experiments are still needed to ensure that incorporation of signal amplification meets the design criteria of MRMBI probes.

Finally, with the development and characterization of an optimized probe design for multiplexed and reiterative molecular biomarker imaging completed, the development of signal amplification schemes and of methods through which these probes can be easily conjugated to ligands (proteins, antibodies, mRNA, aptamers etc) that recognize biomarkers of interest will increase the diagnostic power of individual histological specimens and open the door to more sophisticated analyses of cell phenotype and its functional relationship to disease.

References

1. Dalton, W.S. & Friend, S.H. Cancer biomarkers-an invitation to the table. *Science* 312, 1165-1168 (2006).
2. Sawyers, C.L. The cancer biomarker problem. *Nature* 452, 548-552 (2008).
3. Goulart, B.H.L., Clark, J.W., Pien, H.H., Roberts, T.G., Finkelstein, S.N. & Chabner, B.A. Trends in the use and role of biomarkers in phase I oncology trials. *Clin. Cancer Res.*, 13, 6719-6726 (2007).
4. Lippman, S. M. *et al.* Biomarkers as intermediate end points in chemoprevention trials. *J. Natl. Cancer Inst.* 82, 555-560 (1990).
5. Koomen, J.M. *et al.* Proteomic contributions to personalized cancer care. *Mol. Cell. Proteomics* 7, 1780-1794 (2008).
6. Bild, A.H. *et al.* Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 439, 353-357 (2006).
7. Dowsett, M & Dunbier, A.K. Emerging biomarkers and new understanding of traditional markers in personalized therapy for breast cancer. *CCR Focus* 14, 8019-8016 (2008).
8. Spinny, L. Caught in time. *Nature* 442, 736-738 (2006).
9. Han, Z. *et al.* Noninvasive assessment of cancer response to therapy. *Nat. Med.* 14, 343-349 (2008).
10. Sheehan, K.M. *et al.* Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. *Mol Cell Proteomics* 4, 346-355 (2005).
11. Tibes, R., *et al.* Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. *Mol. Cancer Ther.* 5, 2512-2521 (2006).
12. Daub, H. *et al.* Kinase-selective enrichment enables quantitative phosphoproteomics of the kinome across the cell cycle. *Mol. Cell* 31, 438-448 (2008).
13. Norio Tonotsuka, Yoshio Hosoi, Shukichi Miyazaki, Go Miyata, Ko Sugawara, Takahiro Mori, Noriaki Ouchi, Susumu Satomi, Yoshihisa Matsumoto, Keiichi Nakagawa, Kiyoshi Miyagawa, Tetsuya Ono. *International Journal of Molecular*

Medicine (2006) Heterogeneous expression of DNA-dependent protein kinase in esophageal cancer and normal epithelium 18, 441-447

14. Tae, K. *et al.* Expression of vascular endothelial growth factor and microvessel density in head and neck tumorigenesis. *Clin Cancer Res* 6, 2821-2828 (2000).
15. Tsurui, H., *et al.* Seven-color fluorescence imaging of tissue samples based on fourier spectroscopy and singular value decomposition. *J. Histochem. Cytochem.* 48, 653-662 (2000).
16. Van der Loos, C.M. Multiple immunoenzyme staining: methods and visualizations for the observation with spectral imaging. *J. Histochem. Cytochem.* 56, 313-328 (2008).
17. Brounsa, I., *et al.* Triple immunofluorescence staining with antibodies raised in the same species to study the complex innervation pattern of intrapulmonary chemoreceptors. *J. Histochem. Cytochem.* 50, 575-582 (2002).
18. Mittag Anja, Lenz Dominik, Gerstner Andreas O.H and Tarnok Attila (2006)Hyperchromatic cytometry principles for cytomics using slide based cytometry. *Cytometry.* 69A, 691-703.
19. Lippman, S.M. & Lee, J.J. Reducing the "risk" of chemoprevention: defining and targeting high risk. *Cancer Res.* 66 2893-2903, (2006).
20. Rikova, K. *et al.* Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131, 1190-1203 (2007).
21. Image courtesy of Walter N.Hittelman, Department of Therapeutics, MD Anderson Cancer Center.
22. Al-Hajj Muhammad, Wicha Max S., Benito-Hernandez Adalberto, Morrison Sean J. and Clarke Michael F. (2003) Prospective identification of tumorigenic breast cancer cells. *PNAS.* 100, 3983-3988.
23. Klonisch Thomas, Wiechec Emilia, Hombach-Klonisch Sabine, Ande Sudharsana R., Wesselborg Sebastian, Schulze-Osthoff Klaus and Los Marek. (2008) Cancer stem cell markers in common cancers – therapeutic implications. *Trends in Mol.Med.* 14(10), 450- 460.
24. Woodward Wendy A. and Sulman Erik P. (2008) Cancer stem cells: markers or biomarkers. *Cancer Metastasis Rev*27, 459-470
25. Al-Hajj Muhammad and Clarke Michael F. (2004) Self-renewal and solid tumor stem cells. *Oncogene.* 23, 7274-7282.

26. Hittelman Walter N. and Mila Luka (2008) Cancer Stem Cells and Tumor Response to Therapy: Current Problems and Future Prospects. *Seminars in Radiation Oncology*. 11, 96-105.
27. Li Linheng and Neaves William B. (2006) Normal stem cells and cancer stem cells: The niche matters. *Cancer Res*. 66(9), 4554-4557
28. Ramos-Vara, J.A. Technical aspects of immunohistochemistry. *Vet. Pathol*. 42, 405-426 (2005).
29. BeavisAJ, Pennline KJ.(1996) ALLO-7: A new fluorescent tandem dye for use in flow cytometry. *Cytometry* 24, 390–394.
30. Roederer M, Kantor AB, Parks DR, Herzenberg LA. (1996) Cy7PE and Cy7APC:Bright new probes for immunofluorescence. *Cytometry*. 24, 191–197.
31. Roederer M, De Rosa S, Gerstein S, Anderson M, Bigos M, Stovel R, Nozaki T, Parks D, Herzenberg L, Herzenberg L. (1997). 8-Color, 10-parameter flow cytometry to elucidate complex leukocyte heterogeneity. *Cytometry*. 29, 328–339.
32. Bruchez, M. Jr. *et al.* Semiconductor nanocrystals as fluorescent biological labels. *Science* 281 2013-2016 (1998).
33. Qu, L. & Peng, X. Control of photoluminescence properties of CdSe nanocrystals in growth. *J. Am. Chem. Soc.* 124, 2049–2055 (2002).
34. Xing, Y., *et al.* Bioconjugated quantum dots for multiplexed and quantitative immunohistochemistry. *Nat. Protocols* 2, 1152-1165 (2007).
35. Roger A. Schultz, Thomas Nielsen, Jeff R. Zavaleta, Raynal Ruch, Robert Wyatt, and Harold R. Garner (2001). Hyperspectral Imaging: A novel approach for microscopic analysis. *Cytometry* 43, 239–247.
36. Pirici, D. *et al.* Antibody elution method for multiple immunohistochemistry on primary antibodies raised in the same species and of the same subtype. *J. Histochem. Cytochem.* DOI:953240 (2009).
37. Toth Z.E. & Mezey, E. Simultaneous visualization of multiple antigens with tyramide signal amplification using antibodies from the same species. *J. Histochem. Cytochem.* 55, 545-554 (2007).
38. Suzuki T, Tate G, Ikeda K and Mitsuya T. (2005) A novel multicolor immunofluorescence method using heat treatment. *Acta Med Okayam.* 59(4),145-51.

39. Schubert, W., *et al.* Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. *Nat. Biotechnol.* 24, 1270-1278 (2006).
40. Wiebke Laffers, Anja Mittag, Dominik Lenz, Attila Tarnok and Andreas O.H Gerstner (2006) Iterative restaining as a pivotal tool for n-color immunophenotyping by slide based cytometry. *Cytometry.* 69A, 127-130
41. Stojanovic M. N., Mitchell T. E., Stefanovic D (2002) Deoxyribozyme-based logic gates. *J. Am. Chem. Soc.* 124, 3555–3561
42. Seelig G., Soloveichik D., Zhang D. Y., Winfree E. (2006) Enzyme-free nucleic acid logic circuits. *Science* 314, 1585–1588.
43. Hagiya M., Yaegashi S., Takahashi K. (2006) Computing with hairpins and secondary structures of DNA. *Nanotechnology, science and computation.* 293–308.
44. Dirks, R.M. and Pierce, N.A. (2004) Triggered amplification by hybridization chain reaction. *Proc. Natl. Acad. Sci.* 101, 15275–15278.
45. Daniel Lubrich, Simon J. Green and Andrew J. Turberfield (2009) kinetically controlled self-assembly of DNA oligomers. *J. Am.Chem. Soc.*131, 2422-2423
46. David Yu Zhang and George Seelig. (2011) Dynamic DNA nanotechnology using strand displacement reactions *Nature Chemistry* 3, 103-113
47. Yurke B, Mills AP Jr, Cheng SL. (1999) DNA implementation of addition in which the input strands are separate from the operator strands. *Biosystems* 52, 165-174.
48. David Yu Zhang (2011). Cooperative Hybridization of Oligonucleotides. *J. Am. Chem. Soc.* 133, 1077–1086.
49. Zhang D. Y.; Winfree E. (2006) Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc.* 131, 17303–17314.
50. Zhang D. Y.; Winfree E. (2010) Robustness and modularity properties of a non-covalent DNA catalytic reaction. *Nucleic Acids Res.* 38, 4182–4197
51. Duose D. Y, Schweller R. M, Hittelman W. N and Diehl M. R (2010). Multiplexed and Reiterative Fluorescence Labeling via DNA Circuitry. *Bioconjug. Chem.* 21(12), 2327-2331.
52. Choi H.M.T., Chang J.Y., Trinh L.A., Padilla J.E., Fraser S.E., and Pierce N.A. (2010) Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature Biotechnol* 28, 1208-1212

53. Yaakov Benenson, Binyamin Gil, Uri Ben-Dor, Rivka Adar & Ehud Shapiro (2004) An autonomous molecular computer for logical control of gene expression. *Nature* 429, 423-429
54. Zhang D. Y.; Turberfield A. J.; Yurke B.; Winfree E. (2007) Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* 318, 1121–1125.
55. Zuker M., *Nucleic Acids Res.* (2003) OligoArray 2.0: Thermodynamically improved oligonucleotide design for microarrays. 31 (12), 3057-3062
56. Schweller, R.M. Constantinou, P.E., Frankel, N.W. Narayan, P. & Diehl, M.R. (2008) Design of DNA-Conjugated Polypeptide-based capture probes for the anchoring of protein to DNA matrices. *Bioconj. Chem.* 19, 2304-2307
57. Ishimatsu Uto, Hirayama T., Ueda H., Tsuruta S., and Kambara T. (1991). Determination of urinary tamm-horsfall protein by elisa using a maleimide method for enzyme-antibody conjugation. *J. Immunol. Methods* 138(1), 87-94.
58. Zhang David Yu and Winfree Erik (2009) Control of DNA strand displacement kinetics using toehold exchange *J. Am.Chem. Soc.* 131, 17303-17314
59. Peterson A. W.; Wolf L. K.; Georgiadis R. M. (2002) Hybridization of mismatched or partially matched DNA at surfaces. *J. Am. Chem. Soc.* 124, 14601–14607.
60. Carlon E.; Heim T. (2006) Thermodynamics of RNA/DNA hybridization in high-density oligonucleotide microarrays. *Physica A* 362, 433–449.
61. Zhang David Yu and Winfree Erik (2010) Robustness and modularity properties of a non-covalent DNA catalytic reaction. *Nucleic Acids Research*, 38 (12), 4182-4197
62. Yin P.; Choi H. M. T.; Calvert C. R.; Pierce N. A. (2008) Programming biomolecular self-assembly pathways. *Nature* 451, 318–322.
63. Moll, J. R., Ruvinov, S. B., Pastan, I., and Vinson, C. (2001). Designed heterodimerizing leucine zippers with a range of p*K*s and stabilities up to 10¹⁵ s. *Protein Sci.* 10(3), 649-655.
64. <http://blast.ncbi.nlm.nih.gov/Blast.cgi>
65. <http://www.idtdna.com/>
66. Lilley David M. J. (2000). Structures of helical junctions in nucleic acids. *Quarterly Reviews of Biophysics* 33, 109–159

67. <http://www.nupack.org/partition/new>
68. Feng Bobo, Frykholm Karolin, Norde Bengt and Westerlund Fredrik (2010) DNA strand exchange catalyzed by molecular crowding in PEG solutions Chem. Commun., 46, 8231-8233
69. Karymov Mikhail, Daniel Douglas, Sankey Otto F., Lyubchenko Yuri L. (2005). Holliday junction dynamics and branch migration: Single-molecule analysis. Proc. Natl. Acad. Sci.102 (23), 8186-8191
70. Sloan David J and Hellinga Homme W (1999) Dissection of the protein GB1 domain binding site for human IgG Fc fragment. Protein Science (8), 1643- 1648

Appendix A

Table 1: List of oligonucleotide sequences used in design of DNA-circuits

Underlined oligonucleotides denote toehold domains. All strands are listed in 5' to 3' direction. /3AmMC6/ denotes a 3' amino modifier on carbon 6; /5Cy5/ denotes 5' cy5 fluorophore; /5Cy3/ denotes 5' cy3 fluorophore and /3IAbrQSp/ denotes 3' Iowa black quencher.

Sequence Name	Nucleic Acid Sequence
Circuit1_O1	CCA CAT ACA TCA TAT TCC CTC ATT CAA TAC CCT ACG/3IAbrQSp/
Circuit1_O2	CTT TCC TAC A CC TAC GTC TCC AAC TAA CTT ACG G
Circuit1_LB	<u>TGG AGA</u> /iCy5/CGT AGG GTA TTG AAT <u>GAG GGC</u> CGT AAG TTA GTT GGA GAC GTA GG
Circuit1_C	CAT TCA ATA CCC TAC GTC TCC ATT TTT TTT TT /3AmMC6/
Circuit1_Fuel	CCT ACG TCT CCA ACT AAC TTA CGG CCC TCA TTC AAT ACC CTA CG/3IAbrQSp/
Circuit1_O1 comp	GTA TTG AAT GAG GGA ATA TGA TGT ATG TGG
Circuit1_O1short	CCA CAT ACA TCA TAT TCC CTC ATT
Circuit2_O1	ACC TCT TCA CGA ACA TTT CA/3IAbrQSp/
Circuit2_O2	ACC TAA TAG C AC CAC ATC AAT CTC GAT CCA GTA C
Circuit2_LB	<u>TGG CTA</u> /iCy3/TGA AAT GTT CGT GAA <u>GAG GTG</u> TAC TGG ATC GAG ATT GAT GTG GT
Circuit2_C	CTT CAC GAA CAT TTC ATA GCC ATT TTT TTT TT /3AmMC6/
Circuit2_Fuel	ACC ACA TCA ATC TCG ATC CAG TAC ACC TCT TCA CGA ACA TTT CA
Circuit2_O1comp	TGT TCG TGA AGA GGT
Circuit2_O1short	ACC TCT TCA
Circuit3_O1	TCA CAC ATC AAC CTC T TCTT T CTC TCG ACA CAT CAC
Circuit3_O2	CTT TCC TAC A CT TAT TCA TCC TTT CAC TCA CTT C
Circuit3_LB	<u>GAA GTG</u> AGT GAA AGG ATG AAT <u>AAG AAG</u> AGTG ATG TGT CGA GAG AAAG TAA

Circuit3_C	TCT CTC GAC ACA TCA C TTA CTT TT TTT TTT TT /3AmMC6/
Circuit3_Fuel	CT TAT TCA TCC TTT CAC TCA CTT CTCTT TCT CTC GAC ACA TCA C
Circuit3_O1comp	TGT CGA GAG AAA GAA GAG GTT GAT GTG TGA
Circuit3_O1short	TCA CAC ATC AAC CTC TTC TTT CTC
Circuit4_SB	CAC CAA CCC AAT TCT C CCTA C CCA TTC CTG TAT CAT
Circuit4_OB	ACC TAA TAG C TAC CTT CCC TCT ATT CAT GTC CAC
Circuit4_LB	GTG GAC ATG AAT AGA GGG AAG <u>GTATAG</u> GATG ATA CAG GAA TGG GTGG AGT
Circuit4_C	CCC ATT CCT GTA TCA T AC TCC A TT TTT TTT TT /3AmMC6/
Circuit4_Fuel	TAC CTT CCC TCT ATT CAT GTC CAC CCT ACC CAT TCC TGT ATC AT
Circuit4_O1comp	CAG GAA TGG GTA GGG AGA ATT GGG TTG GTG
Circuit4_O1short	CAC CAA CCC AAT TCT C CCTA C CCA
Circuit5_O1	CAT ACC ACA ACA ATT TAC TTC ACC AAC CCA TCC ACT
Circuit5_O2	CTT TCC TAC AAA TCG CCA AAC TAC AAA CTC AAT C
Circuit5_LB	<u>GAG GTG</u> AGT GGA TGG GTT GGT <u>GAA GTG</u> ATT GAG TTT GTA GTT TGG CGA TT
Circuit5_C	CAC CAA CCC ATC CAC TCA CCT C TT TTT TTT TT /3AmMC6/
Circuit5_Fuel	AAT CGC CAA ACT ACA AAC TCA ATC ACTT CAC CAA CCC ATC CAC T
Circuit5_O1comp	TGG GTT GGT GAA GTA AAT TGT TGT GGT ATG
Circuit5_O1short	CAT ACC ACA ACA ATT TAC TTC ACC
1compcy5	/5Cy5/TGT AGG AAA G

