ISOLATION OF HOMOGENOUS CARDIAC CELL POPULATIONS FROM DIFFERENTIATING PLURIPOTENT STEM CELLS USING MOLECULAR BEACONS

A Dissertation
Presented to
The Academic Faculty

by

Brian Wile

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Engineering

Georgia Institute of Technology
May 2014

ISOLATION OF HOMOGENOUS CARDIAC CELL POPULATIONS FROM DIFFERENTIATING PLURIPOTENT STEM CELLS USING MOLECULAR BEACONS

Approved by:

Dr. Gang Bao, Ph.D., Advisor
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Dr. Chunhui Xu, Ph.D.
Department of Pediatrics
Emory University and Children's Hospital of Atlanta

Dr. W. Robert Taylor, MD, Ph.D
Department of Cardiology
Emory University, Georgia Institute of Technology and Atlanta VA

Dr. Charlie Searles, MD.
Department of Cardiology
Emory University and Atlanta Veterans Affairs Medical Center

Dr. Phil Santangelo, Ph.D
Deparment of Biomedical Engineering
Georgia Institute of Technology and Emory University

Date Approved: March 20th, 2014

ACKNOWLEDGEMENTS

I would like to thank my lab mates and collaborators, who have taught me more than I could write in even 1,000 pages. I'm also thankful for my undergraduate research assistants who helped to speed up my research and were a pleasure to work with. I would also like to thank my advisor who provided me with space, funding, and support throughout my graduate studies.

I would like to acknowledge my mother and father, who have always supported my education, and my friends, who have always put up with research schedules. I would most like to thank my girlfriend Jennifer and my dog Woodford, who have kept me grounded and focused on the important things in life even when things get stressful.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF SYMBOLS AND ABBREVIATIONS

RNA	Ribonucleic Acid
mRNA	Messenger Ribonucleic Acid
DNA	Deoxyribonucleic acid
ODN	Oligonucleotide
CM	Cardiomyocyte
MB	Molecular Beacon
PCR	Polymerase Chain Reaction
FISH	Fluorescence In Situ Hybridization
RNP	Ribonucleoprotein
SNP	Single nucleotide polymorphism
CPP	Cell-penetrating peptide
SLO	Streptolysin O
GFP	Green fluorescent protein
SAGE	Serial analysis of gene expression
MHC	Myosin Heavy Chain
TNT	Troponin T
AP	Action potentials
APD	Action potential duration
CF	Cardiac fibroblasts
CHD	Congenital heart defects
CPP	Cell-penetrating peptides
EB	Embryoid body
ECC	Embryonic carcinoma cell
ESC	Embryonic stem cells
FACS	Fluorescent activated cell sorting
H&E	Hematoxylin and eosin
LNA	Locked nucleic acid
MB	Molecular beacons
MEC	Mouse aortic endothelial cells
MI	Myocardial infarction
NLS	Nuclear localization signal
NPPA	Natriuretic Peptide A
PI	Post-infection
RNP	Ribonucleoprotein
SMC	Smooth muscle cells
SNP	Single nucleotide polymorphisms
CHO	Chinese Hamster Ovary
CT	Threshold Cycle
DAPI	4',6-Diamidino-2-Phenylindole
FAM	Fluorescein Amidite
IgG	Immunoglobulin G
ISH	In Situ Hybridization
MYOD	Myogenic Differentiation 1
NEUROD	Neurogenic Differentiation 1

NIR	Near Infrared	PSC	Pluripotent Stem Cell
PBS	Phosphate Buffered Saline	STO	Mouse Embryonic Fibroblasts
PNA	Peptide Nucleic Acid	VCM	Ventricular Cardiomyocyte
FRET	Fluorescence Resonance Energy Transfer		
BLAST	Basic Local Alignment Search Tool		
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats		
TALEN	Transcription Activator-like Effector Nucleases		
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase		
ALCADIA	AutoLogous Human. CArdiac/Derived Stem Cell To Treat. Ischemic cArdiomyopathy		
CADUCEUS	CArdiosphere-Derived aUtologous stem CElls to reverse ventricULar dySfunction		
SCIPIO	Stem Cell Infusion in Patients with Ischemic Cardiomyopathy		
TMRM	Tetramethylrhodamine, methyl ester		

SUMMARY

Human pluripotent stem cells (hPSCs) hold the potential to revolutionize cardiac tissue engineering. Because of their ability to proliferate and differentiate into all cardiomyocyte subtypes they represent an opportunity to regenerate virtually any tissue lost from the over 1 million cardiac disease patients in the United States alone. Studies have shown, however, that hPSCs which are not terminally differentiated pose a variety of risks including teratoma formation and lack of appropriate cell engraftment. It is therefore important to ensure that only well characterized cardiac subtypes are implanted into patients or used for research purposes. Current differentiation protocols generate a mixture of cardiac subtypes, and research on cardiac subtype specification is hampered by the lack of a high throughput method to distinguish cardiac subtypes.

This thesis establishes the ability to identify, enrich and characterize cardiac subtypes using MBs. This will provide a robust tool for clinical use of hPSCs in cardiac cell therapy and for analysis of differentiation protocol effects on cardiac subtype formation.

CHAPTER 1

MOLECULAR BEACONS: TOOLS TO DETECT MESSENGER RNA

Molecular Beacon description and function

The ability to image specific ribonucleic acid (RNA) in living cells in real time can provide essential information on RNA synthesis, processing, transport, and localization. Visualizing the dynamics of RNA expression and localization in response to external stimuli will offer unprecedented opportunities for advancement in molecular biology, disease pathophysiology, drug discovery, and medical diagnostics. Over the past decade evidence has revealed the essential role that RNA molecules have in a wide range of functions in living cells, from physically conveying and interpreting genetic information, to essential catalytic roles, to providing structural support for molecular

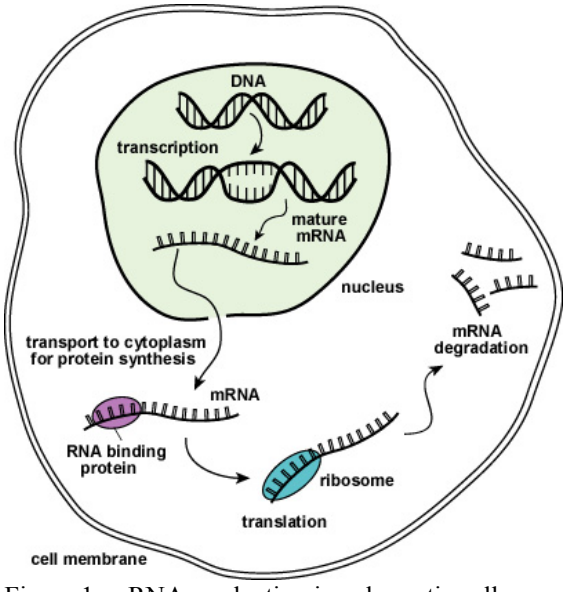

Figure 1, mRNA production in eukaryotic cells.

machines, and to modifying gene expression (Figure 1). These functions are realized through control of the expression level and stability, both temporally and spatially, of specific RNAs in a cell. Therefore, determining the dynamics and localization of RNA molecules in living cells will significantly impact molecular biology and medicine.

Of particular interest is the fluorescent imaging of specific messenger RNAs (mRNAs), both their expression level and subcellular localization, in living cells. As shown schematically in Figure 1, for eukaryotic cells a pre-mRNA molecule is synthesized in the cell nucleus. After processing, such as splicing and polyadenylation, the mature mRNAs are transported from the cell nucleus to specific sites in the cytoplasm. The mRNAs are then translated by ribosomes before being degraded by Rnases. The limited lifetime of mRNA enables a cell to alter protein synthesis rapidly in response to its changing needs. mRNA is always complexed with RNA-binding proteins to form a ribonucleoprotein (RNP). This has significant implications to the live cell imaging of mRNAs. Many in vitro methods have been developed to provide a relative (mostly semiquantitative) measure of gene expression level within a cell population using purified DNA or RNA obtained from cell lysate. These methods include polymerase chain reaction (PCR)(*4*), Northern hybridization (or Northern blotting) (*5*), expressed sequence tag (EST) (*6*), serial analysis of gene expression (SAGE) (*7*), differential display (*8*), and DNA microarrays (*9*). These technologies, combined with the rapidly increasing availability of genomic data for numerous biological entities, present exciting possibilities for understanding human health and disease. For example, pathogenic and carcinogenic sequences are increasingly being used as clinical markers for diseased

states. However, using in vitro methods to detect and identify foreign or mutated nucleic acids is often difficult in a clinical setting due to the low abundance of diseased cells in blood, sputum, and stool samples. Further, these methods cannot reveal the spatial and temporal variation of RNA within a single cell.

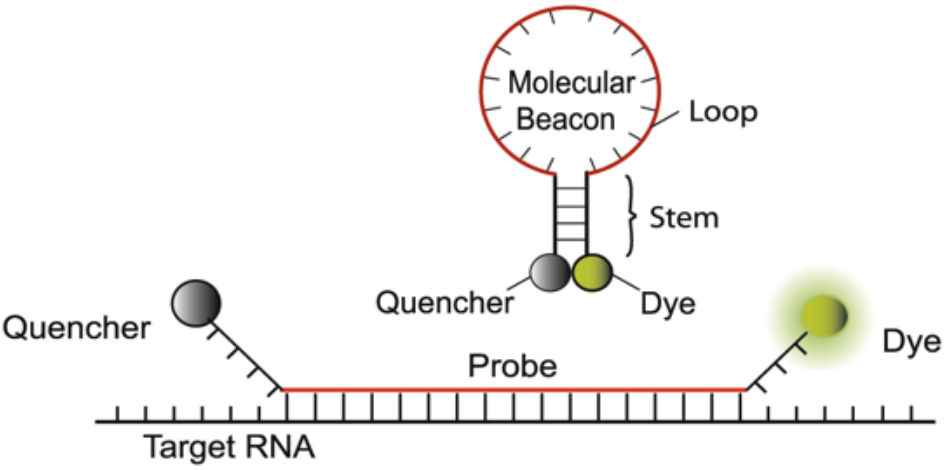

Figure 2, Illustration of the structure of a molecular beacon.

Labeled linear oligonucleotide (ODN) probes have been used to study intracellularmRNA via in situ hybridization (ISH)(*10*), in which cells are fixed and permeabilized to increase the probe delivery efficiency. Unbound probes are removed by washing to reduce background and achieve specificity(*11*). To enhance the signal level, multiple probes targeting the same mRNA can be used (*10*). However, fixation agents and associated chemicals can have a considerable effect on signal level (*12*) and the

integrity of certain organelles, such as mitochondria. Fixation of cells by either cross-linking or denaturing agents and the use of proteases in ISH assays may prevent the obtaining of an accurate description of intracellular mRNA localization. It is also difficult to obtain a dynamic picture of gene expression in cells using ISH methods. In order to detect RNA molecules in living cells with high specificity, sensitivity, and signal-to-background ratio, especially for low abundance genes and clinical samples containing a small number of diseased cells, the probes need to recognize RNA targets with high specificity, convert target recognition directly into a measurable signal, and differentiate between true and false positive signals. It is important for the probes to quantify low gene expression levels with high accuracy, and have fast kinetics in tracking alterations in gene expression in real time. For detecting genetic alterations such as mutations, insertions, and deletions, the ability to recognize single nucleotide polymorphisms (SNPs) is essential. To achieve this optimal performance, it is necessary to have a good understanding of the structure–function relationship of the probes, probe stability, and RNA target accessibility in living cells. It is also necessary to achieve efficient cellular delivery of probes.

MBs are a class of ODN probes that have been used previously for live cell RNA imaging. As illustrated in Figure 2, these probes are labeled at one end with a reporter fluorophore and at the other end with a quencher. MBs are designed to form a stem-loop hairpin structure in the absence of a complementary target so that fluorescence of the fluorophore is quenched. Hybridization with the target nucleic acid opens the hairpin and physically separates the fluorophore from quencher, allowing a fluorescence signal to be emitted upon excitation (Figure 2). This enables a molecular beacon to function as a

sensitive probe with a high signal-to-background ratio. Under optimal conditions, the fluorescence intensity of MBs can increase by > 200-fold upon binding to their targets. The ability to transduce target recognition directly into a fluorescence signal with high signal-to-background ratio has allowed MBs to enjoy a wide range of biological and biomedical applications, including real time monitoring of PCR, genotyping and mutation detection, multiple analyte detection, assaying for nucleic acid cleavage in real time, cancer cell detection, studying viral infection, and monitoring RNA expression and localization in living cells. MBs provide an enticing opportunity to optically track some of the most elusive events in molecular biology. Despite the complexity of interpreting molecular beacon results, MBs themselves are relatively simple molecules to synthesize. Designing MBs involves understanding a series of trade-offs between the specificity, sensitivity, and signal-to-background for beacons. In the following sections we discuss how the major aspects of a molecular beacon (probe sequence, hairpin structure, and fluorophore/quencher selection) determine beacon performance.

Principle factors affecting the design of MBs

Specificity

The most important property of a molecular beacon is its specificity, which can be loosely quantified as the difference in melting temperature between perfectly complementary hybrids and hybrids with single base mismatches. MBs typically exhibit a higher specificity for perfectly complementary nucleic acid targets than linear ODN probes do. It has been shown that properly designed MBs can readily discriminate between targets that differ by as little as a single nucleotide (*13, 14*). The reason is the energy penalty associated with unwinding the molecular beacon stem reduces binding to

mismatched targets because it is less energetically favorable. In experiments where the detection of SNPs is required, the specificity of MBs can be improved by increasing the stem length. A longer stem provides a wider set of conditions over which MBs can discriminate between two targets. This can be attributed to the enhanced stability of the molecular beacon stem–loop structure and the resulting smaller free energy difference between closed (unbound) MBs and molecular beacon–target duplexes, which generates

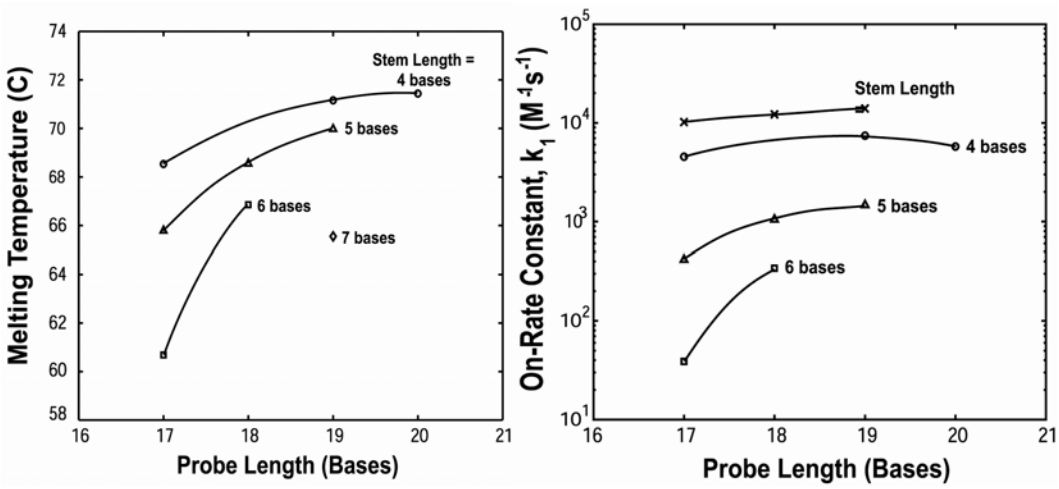

Figure 3, The effect of stem and loop length on the melting temperature of MBs.
A) Melting temperature of MBs based on loop and stem length. B) Frequency of binding ebents based on identical loop and stem length conditions.

a condition where a single base mismatch reduces the energetic preference of probe–target binding. A longer stem also increases the signal-to-background ratio, since the more stable hairpin conformation reduces the probability of stem opening due to Brownian fluctuations and results in more efficient quenching of the fluorescent dye. The

type and number of nucleotides in a molecular beacon's loop must be carefully designed for the chosen RNA target and the set of experimental conditions (*13, 15*). For example, in live cell studies it is desirable for the melting temperature of perfectly complementary hybrids to be above 37 ◦C and the melting temperature for hybrids with single base mismatches to be less than 37 ◦C (Figure 3). Both of these melting temperatures will increase with increasing loop length and decreasing stem length. If the melting temperatures are too high, it would not be possible to discriminate between perfectly complementary and mismatched targets under physiological conditions. On the other hand, if the melting temperatures are too low, a large number of MBs may open under physiological conditions, leading to a high level of background signal. The melting temperature of a molecular beacon can be tailored by changing its stem–loop structure, as demonstrated in Figure 3A. Changing the probe length of a molecular beacon may also influence the rate of hybridization, as demonstrated by Figure 3B, but generally to a lesser extent than changing the stem length. While both the stability of the hairpin probe and its ability to discriminate targets over a wider range of temperatures increase with increasing stem length, this is accompanied by a decrease in hybridization on-rate constant, as shown in Figure 3B. For example, MBs with a four-base stem had an on-rate constant up to 100 times greater than MBs with a six-base stem.

In addition to melting temperature considerations, the characteristics of the target sequence itself must be unique to guarantee specificity. The Basic Local Alignment Search Tool (BLAST) developed by the National Center for Biotechnology Information (NCBI) (*16*) or similar software can be used to select multiple target sequences of 15–25 bases that are unique for the target RNA. For any target sequence selected there might be

multiple genes in the mammalian genome that have sequences differing by only a few bases. BLAST allows the selection of beacon target sequences that are not only unique for the specific RNA of interest, but also with minimal partial match between beacon sequence and the sequence of other RNAs. It is also necessary to select a target sequence with a balanced G–C content (the percentage of G–C pairs). A G–C mismatch carries a much larger energy penalty, enhancing specificity; however, the prevalence of C–G islands in the genome makes beacons with a high G–C content more likely to have off-target interactions. Therefore, it is important to select a G–C-balanced and unique target sequence to ensure the specificity of MBs.

Several approaches can be taken to validate the signal specificity in live cells. The most common approach is to up- or down-regulate the expression level of a specific RNA and compare molecular beacon based imaging in live cells with RT-PCR data quantifying the RNA expression level. Complications may arise when the approach used to change the RNA expression level in living cells has an effect on multiple genes, leading to some ambiguity even when PCR and beacon results match. A common way to down-regulate the level of a single mRNA in live cells is to use small interfering RNA (siRNA) treatment, which typically leads to >80% reduction of the specific mRNA level when the siRNA protocol is properly optimized.

Sensitivity

Once specificity has been guaranteed, the beacon must also be optimized for sensitivity, eg, maximizing the change in fluorescent intensity based upon a given concentration of target mRNA. Selecting an appropriate dye–quencher pair yields important benefits for the signal to background ratio, multiplexing ability, and

fluorescence quantification of MBs. For example, a systematic study on a wide range of fluorophore–quencher combinations showed that the quenching efficiency (contact quenching) could vary between 57% and 98%(*17*). Quenching efficiencies could potentially be further improved by either using inorganic quenchers, such as gold, or by incorporating multiple quenchers into a single molecular beacon (*18, 19*). Alternatively, the signal intensity of MBs can be increased by using quantum dots or photoluminescent polymers (*20-22*). Initial studies have suggested that quantum-dot-based MBs only exhibit a signal to background ratio of 6 : 1 due to inefficient quenching. Inefficient quenching is unlikely to be an issue for photoluminescent polymers, which exhibit a superquenching effect. When the fluorescence of any single repeat unit is quenched, the entire polymer chain responds in the same fashion. Long polymer chains can, therefore, be used to provide an amplified fluorescent signal that can be modulated by a single quencher. However, nonspecific interaction of the photoluminescent polymers prevents them from widespread application in live cell RNA detection assays (*22*). A more conventional way to increase signal-to-background ratio is to use multiple beacons to target the same RNA molecule. As an example, MBs were designed to target a sequence in the genome of bovine respiratory syncytial virus (bRSV) that has three exact repeats (*23*). Figure 4 shows the molecular beacon signal indicating the spreading of viral infection at days one, three, five, and seven post-infection (PI), which demonstrates the ability of MBs to monitor and quantify in real time the viral infection process. MBs were further used to image the viral genomic RNA (vRNA) of human RSV (hRSV) in live Vero cells, revealing the dynamics of filamentous virion egress, and providing insight into how viral filaments bud from the plasma membrane of the host cell(*24*).

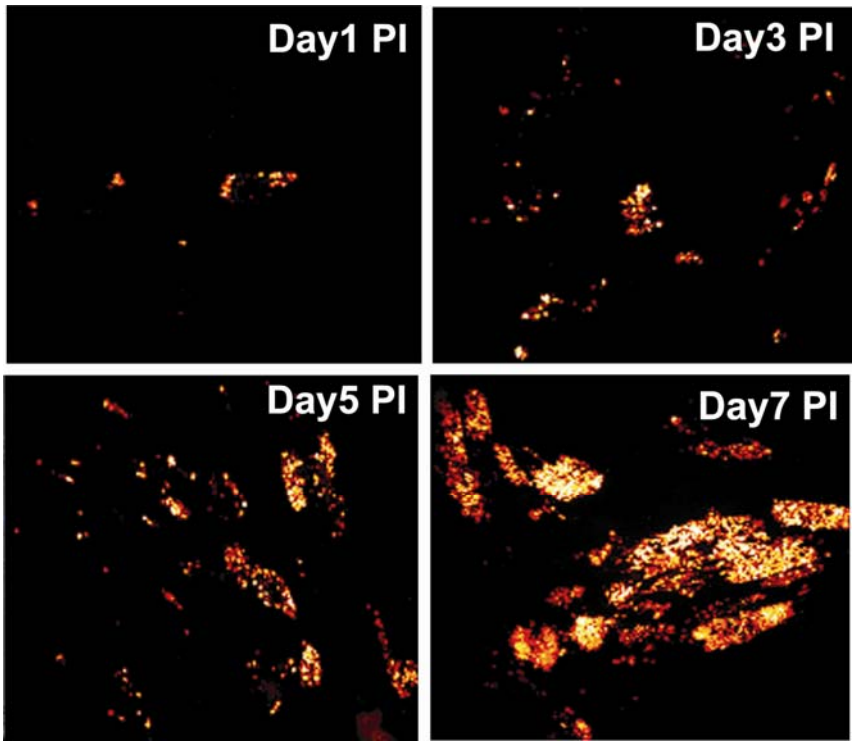

Figure 4, Increasing molecular beacon signal after viral RNA expression.

In addition to creating brighter fluorescent labels or adding multiple labels to each mRNA, the excitation and emission peaks of the labels are also important. Red-shifted fluorophores can be used to improve signal-to-background in live cells by eliminating the interfering effects that result from autofluorescence. It is also possible to use lanthanide chelate as the donor in a dual-FRET probe assay and perform time-resolved measurements to dramatically increase the signal-to-background ratio (*25*). As fluorescent imaging strategies have advanced, there has been a general trend towards more quantitative measurements of fluorescent signals, with the ultimate goal being absolute quantification. The absolute quantification of fluorescence could allow the exact

number of fluorophores within a compartment/cell to be quantified and correspondingly allow the number of target genes, proteins, or enzymes to be quantified. However, factors such as nonspecific protein interactions and pH could have a dramatic effect on the fluorescence intensity of some fluorophores. Therefore, if accurate fluorescence measurements are desirable, it is necessary to select fluorescent labels that are insensitive to their environment. Recently, it has been shown that although the fluorescence intensity of a few fluorophores (e.g., fluorescein) was highly susceptible to the intracellular environment, other fluorophores (e.g., Dylight 649, Alexa647, and Alexa750) were insensitive to the intracellular environment (*26*).

Cellular factors affecting molecular beacon performance

Accessibility

A critical issue in molecular beacon design is target accessibility. It is well known that proteins are constantly bound to functional mRNA molecules in living cells, forming a ribonucleoprotein (RNP). Furthermore, an mRNA molecule often has double stranded portions and forms secondary (folded) structures (**Figure 5**). It is therefore necessary to avoid targeting mRNA sequences that are double stranded or occupied by RNA-binding proteins. These sites require the molecular beacon to compete off the RNA-binding protein or RNA strand in order to hybridize to the target. This competition is hypothesized to contribute to a lack of signal when certain MBs designed for a specific mRNA are delivered to living cells. Although predictions of mRNA secondary structure can be made using software such as Beacon Designer (www.premierbiosoft.com) and mfold (http://www.bioinfo.rpi.edu/applications/mfold/old/dna/), they may be inaccurate due to limitations of the biophysical models used and a limited understanding of protein–

RNA interaction. For each target mRNA, therefore, it may be necessary to synthesize multiple unique molecular beacon sequences along the target RNA and test them in living cells to select the best target sequence.

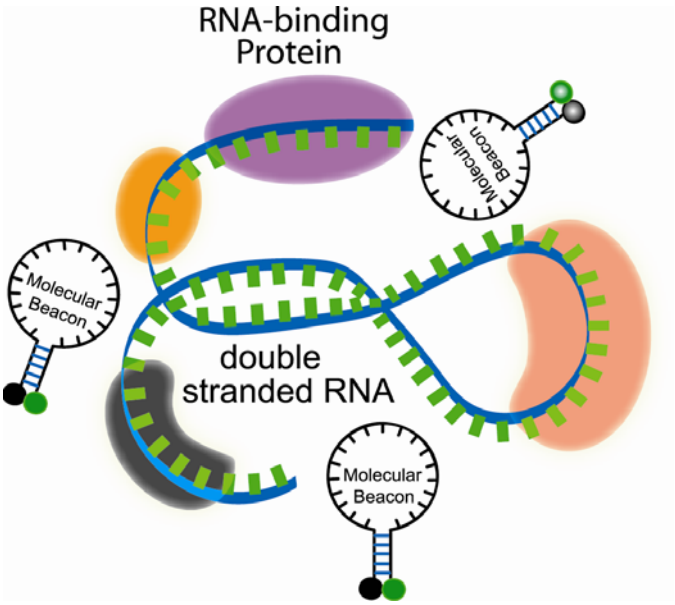

Figure 5, RNA accessibility issues.
The cartoon depicts a target RNA which is bound to RNA-binding proteins and has double stranded regions, preventing the hybridization of molecular beacons.

To uncover the possible molecular beacon design rules, the accessibility of BMP-4 mRNA was studied using different beacon targets for the same mRNA sequence (*27*). Specifically, MBs were designed to target the start codon, the termination codon, siRNA sites, antisense ODN probe sites, and random sites. All the target sequences were run

through the BLAST database to ensure that they were unique to BMP-4 mRNA. Of the eight MBs designed to target BMP-4 mRNA, it was found that only two beacons resulted in a strong signal inside cells. The positive beacons targeted the start codon region and the termination codon region. It was also found that shifting the target sequence of a molecular beacon by just a few bases towards the 3or 5ends would significantly reduce the fluorescence signal from beacons in a live cell assay. This indicates that the target accessibility is quite sensitive to the location of the targeted sequence. These results, together with MBs validated previously, suggest that the start codon and termination codon regions and the exon–exon junctions are often more accessible than other locations in an mRNA.

Delivery of MBs to cells

One of the critical steps in the accurate detection of RNA molecules in living cells is the efficient delivery of synthetic probes into the cytoplasm. ODN-based probes are generally prevented from gaining access to the cytoplasm due to the cell membrane (*28*). Once the probes enter the cells successfully the fraction of probes that are free to hybridize intracellular RNA also becomes a concern. Numerous studies have shown that MBs are rapidly sequestered into the nucleus once introduced into cells; however, there have been an equal number of studies that did not observe this pattern of intracellular distribution. It is not clear whether these differences are cell line dependent, dependent on the method of delivering MBs, or due to some other variable. Nonetheless, several methods have been introduced to prevent nuclear sequestration. Specifically, MBs have been conjugated to large proteins (or nanoparticles) that prevent their passage through nuclear pores and they have been linked to tRNAs that drive nuclear export (*29-32*).

Common techniques that have been used to deliver ODNbased RNA imaging probes into live cells include microinjection, polycationic molecules (such as liposomes and dendrimers), electroporation, cell-penetrating peptides (CPPs) or steptolysin O (SLO).

Microinjection

Microinjection is, perhaps, one of the most direct methods for ensuring the delivery of ODN probes into the cytoplasm of live cells. This method is advantageous because it removes the question of whether beacons have crossed the cell membrane or not. Numerous studies have delivered MBs to a small number of cells and used microscopy to determine the subcellular localization of the probes (*33*). Microinjection is unsuitable for population-based confirmation of beacon results due to the relatively low number of cells that can be analyzed at any given time. Further, microinjection can often be damaging to the cell and may interfere with normal cell function. Because of these drawbacks, many investigators use alternative methods that result in higher delivery throughput and less physical damage than microinjection.

Cationic Transfection Agents

Cationic transfection agents can be used to deliver MBs into living cells. These transfection agents generally form lipoplexes with MBs, which can sometimes stabilize MBs in their hairpin confirmation. While these agents have been found to be effective in some studies (*34*), others have found that many commercial transfection agents result in punctate fluorescent patterns that appear to be indicative of endosomal entrapment (*35*). ODN probes that enter into the endosomal/lysosomal pathway are rapidly degraded by nucleases and the acidic environment (*36*). Consequently, even when transfection methods allow for endosomal/lysosomal escape, only 0.01–10% of the probes retain their

functionality (*37*). Furthermore, probe delivery via the endocytic pathway typically takes 2–4 h. This long time period required increases the likelihood of molecular beacon degradation due to cellular nucleases.

Electroporation

To avoid the deleterious effects associated with endosomal entrapment, methods such as electroporation have been used to deliver ODNs directly into the cytoplasm of living cells. Although, electroporation has traditionally been associated with low cell viability, recent advances in electroporation technology, such as the ability to perform electroporation in microliter-volume spaces (e.g., pipette tips), has led to a reduction in the many harmful events associated with this process, including heat generation, metal ion dissolution, pH variation, and oxide formation. This microliter-volume electroporation process is known as microporation. Recently, it was shown that microporation could lead to the uniform cytosolic distribution of ODN probes in live cells with a transfection efficiency of 93% and an average viability of 86% (*31*). A unique advantage of microporation is that delivery of ODN probes takes seconds, so cells can be analyzed immediately for RNA content. Conversely, a potential disadvantage is that most electroporation techniques require cells to be detached from the culture surface. Therefore, it is several hours before the cells re-adhere to cell culture plate surfaces and RNA localization can be imaged.

Chemical Permeabilization

Another non-endocytic delivery method is toxin-based cell membrane permeabilization. One popular reagent is SLO, which is a pore-forming bacterial toxin that has been used as a simple and rapid means of introducing ODNs into eukaryotic cells

(*38-41*). SLO binds as a monomer to cholesterol and oligomerizes into a ring-shaped structure to form pores of approximately 25–30 nm in diameter, allowing the influx of both ions and macromolecules. An essential feature of this technique is that the toxin-based permeabilization is reversible (*38*). This can be achieved by introducing ODNs with SLO under serum-free conditions and then removing the mixture and adding normal media with serum and calcium (*38, 41*). Since cholesterol composition varies with cell types, the permeabilization protocol needs to be optimized for each cell type by varying temperature, incubation time, cell number, and SLO concentration. Typically, RNA localization can be assessed 30 min to 2 h following the introduction of ODN probes into cells using SLO-based delivery.

Cell-Penetrating Peptide

CPPs have been used to introduce proteins, nucleic acids, and other biomolecules into living cells (*42-44*). Among the family of peptides with membrane translocating activity are antennapedia, HSV-1 VP22, and the HIV-1 Tat peptide. To date, the most widely used peptides are the HIV-1 Tat peptide and its derivatives, owing to their small size and high delivery efficiency. The Tat peptide is rich in cationic amino acids, especially arginines, which are very common in many of the CPPs. However, the exact mechanism for CPP-induced membrane translocation remains elusive. A wide variety of cargos have been delivered to living cells both in cell culture and in tissue using CPPs (*44, 45*). For example, Allinquant et al. (*46*) linked antennapedia peptide to the 5’ end of DNA ODNs (with biotin on the 3 end) and incubated both peptide-linked ODNs and ODNs alone with cells. By detecting biotin using streptavidin–alkaline phosphatase amplification, it was found that the peptide-linked ODNs were internalized very

efficiently into all cell compartments compared with control ODNs. No indication of endocytosis was found. Similar results were obtained by Troy et al.(*47*), with a 100-fold increase in antisense delivery efficiency when ODNs were linked to antennapedia peptides. Peptide-linked MBs were internalized into living cells within 30 min with nearly 100% efficiency. Peptide-based delivery did not interfere with molecular beacon binding to survivin or GAPDH, since similar levels of fluorescence and a similar pattern of localization was seen in cells with beacons delivered using alternative means.

Peptide-linked MBs show impressive potential as an all-in-one molecule capable of self-delivery, targeting, and reporting in live cells. Peptide-linked MBs can also be delivered using an SLO-based approach to target RNA molecules in the cell nucleus by attaching a nuclear localization signal (NLS) peptide to a molecular beacon. MBs designed to target snRNAs U1 and U2 were linked to NLS peptides and delivered to cells using the SLO-based reversible membrane permeabilization method. The small nucleolar RNA U3 was delivered into the nuclei of live HeLa cells, and the localization and co-localization (U1 and U2) of these nuclear RNAs was imaged (*48*). This delivery method can potentially be used to image transcriptional and post-transcriptional processing of RNAs in the nucleus of living cells.

Conclusions and Future Directions

Nanostructured molecular probes such as MBs have the potential to address a wide range of applications that require sensitive detection of genomic sequences. For example, MBs are used as a tool for the detection of single-stranded nucleic acids in homogeneous in vitro assays (*49, 50*). Surface-immobilized MBs used in microarray assays allow for the high throughput parallel detection of nucleic acid targets while avoiding the difficulties associated with PCR-based labeling (*49, 51*). Another novel

application of MBs is the detection of double stranded DNA targets using protein nucleic acids (PNA) "openers" that form triplexes with the DNA strands (*52*). Further, proteins can be detected by synthesizing "aptamer molecular beacon," which, upon binding to a protein, undergoes a conformational change that results in the restoration of fluorescence (*53, 54*).

The most exciting application of nanostructured ODN probes, however, is living-cell gene expression detection. MBs can detect endogenous mRNA in living cells with high specificity, sensitivity, and signal-to-background ratio. Thus, MBs have the potential to provide a powerful tool for laboratory and clinical studies of gene expression in vivo. For example, MBs can be used in high throughput cell-based assays to quantify and monitor the dose-dependent changes of specific mRNA expression in response to different drug leads. The ability of MBs to detect and quantify the expression of specific genes in living cells will also facilitate disease studies, such as viral infection detection and cancer diagnosis.

There are a number of challenges in detecting and quantifying RNA expression in living cells. In addition to issues of probe design and target accessibility, quantifying gene expression in living cells in terms of mRNA copy number per cell poses a significant challenge. It is necessary to distinguish true signal from background signals, to determine the fraction of mRNA molecules hybridized with probes, and to quantify the possible selfquenching effect of the reporter, especially when mRNA is highly localized. Since the fluorescence intensity of the reporter may be altered by the intracellular environment, it is also necessary to create an internal control by, for example, injecting a known quantity of additional fluorescently labeled ODNs into the same cells and obtaining the corresponding fluorescence intensity. Further, unlike in RT-PCR studies, where the mRNA expression is averaged over a large number of cells (usually over 1 million), in optical imaging of mRNA expression in living cells only a relatively small number of cells (typically less than 1000) are observed. Therefore, the average copy

number per cell may change with the total number of cells observed due to the (often large) cell-to-cell variation of mRNA expression.

Another issue in living-cell gene detection using hairpin ODN probes is the possible effect of probes on normal cell function, such as protein expression. As revealed in the antisense therapy research, complementary pairing of a short segment of an exogenous ODN to mRNA can have a profound impact on protein expression levels and even cell fate. For example, tight binding of the probe to the translation start site may block mRNA translation. Binding of a DNA probe to mRNA can also trigger RNase H-mediated mRNA degradation. However, the probability of eliciting antisense effects with hairpin probes may be very low, because low concentrations of probes (<200 nm) are used for mRNA detection in contrast to the high concentrations (typically 20 μm;(*40*)) employed in antisense experiments. Furthermore, antisense effects are generally unable to be observed for at least 4 h, whereas visualization of mRNA with hairpin probes requires less than 2 h after delivery. Well-designed experiments should carry out a systematic study of possible antisense effects, especially for MBs with 2-O-methyl backbone, which may bind to mRNA for longer periods of time. As a new approach for in vivo gene detection, the nanostructured probes can be further developed to have enhanced sensitivity and a wider range of applications. Hairpin ODN probes with near infrared (NIR) dye as the reporter combined with peptide-based delivery have the potential to detect specific RNAs in tissue samples, animals, or even humans. It is also possible to use lanthanide chelate as the donor in a dual-FRET probe assay and perform time-resolved measurements to dramatically increase the signal to noise ratio, thus achieving high sensitivity in detecting low abundance genes.

Although very challenging, the development of these and other nanostructured ODN probes will significantly enhance our ability to image, track, and quantify gene expression in vivo, and provide a powerful tool for basic and clinical studies of human health and disease. There are many possibilities for nanostructured ODN probes to

become clinical tools for disease detection and diagnosis. For example, MBs could be used to perform cell-based early cancer detection using clinical samples, including blood, saliva, and other bodily fluid. In this case, cells in the clinical sample are separated and MBs designed to target specific cancer genes are delivered to the cytoplasm for detecting mRNAs of the cancer biomarker genes. Cancer cells having a high level of the target mRNAs (such as survivin) or mRNAs with specific mutations that cause cancer (such as K-ras codon 12 mutations) would show a high level of fluorescence signal, while normal cells would show just low background signal. In this approach, the target mRNAs would not be diluted compared with the approaches using cell lysate. Thus, a molecular-beaconbased assay has the potential to positively identify cancer cells in a clinical sample with high specificity and sensitivity. It may also be possible to detect cancer cells in vivo by using NIR-dye-labeled MBs in combination with endoscopy. Although there remain significant challenges, imaging methods using nanostructured probes have a great potential in becoming a powerful clinical tool for disease detection and diagnosis.

CHAPTER 2

CARDIOMYOCYTE DEVELOPMENT IN STEM CELLS

Cellular therapies have the potential to replace or regenerate damaged cardiac muscle, a novel therapeutic option that would help millions of patients a year in the US alone (*55*). Previous studies have shown that undifferentiated cells are unsuitable for implantation due to the danger of teratoma formation(*56*), but several experiments and clinical trials have shown that the delivery of terminally differentiated cells is likely to be safe and the next round of studies to demonstrate efficacy is now underway (*57*). Early results purport to show that there is an effect on the injured tissue, although cellular engraftment may not be as important as paracrine signaling effects (*57*). While methodologies for generating cardiomyocytes (CMs) from hPSCs are rapidly improving, numerous hurdles stand in the way of their clinical use. All protocols which do not require genetic manipulation result in heterogeneous populations of cells which may not be as efficacious or safe as terminally differentiated cardiomyocytes (*58-60*). The current gold standard activinA/BMP-4 protocol yielded > 30% of CMs (*61*). Novel methods promise to deliver significantly higher yields of CMs with more than 40% of cells becoming CMs, however none of these techniques can match the purity of genetic manipulation techniques which can lead to up to 98% pure populations. One of the major challenges confronting the field is therefore the development of reproducible isolation techniques that allow scalable purification of cardiomyocytes. Considerable upscaling, effective enrichment, and purification methods should be developed for robust research and clinical use in the future.

Developmental cardiomyocyte biology

Embryonic stem cells can be induced towards the mesodermal pathway by the stimulation of the bone morphogenic protein (BMP) pathway (*62*). Afterwards mesodermal cells begin to form the circulatory cells, and most importantly the heart tube (*63*). Cardiomyocytes (CMs) develop *in vivo* from the mesodermal lineage (Figure 6A), specifically from the lateral plate (*1*). Portions of this tube continue to proliferate, expanding to become recognizable heart chambers, while non-proliferating sections become the conduction system of the heart (Figure 6B)(*1*).

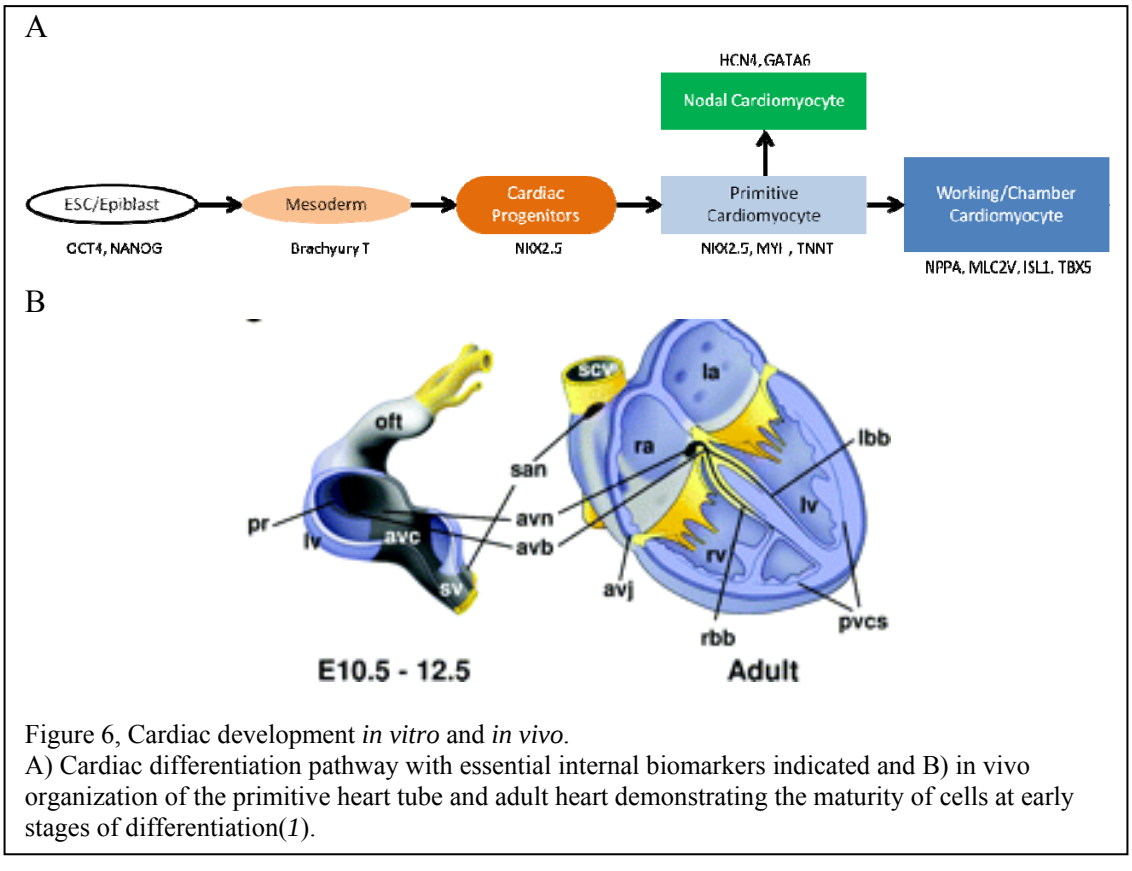

Figure 6, Cardiac development *in vitro* and *in vivo*.
A) Cardiac differentiation pathway with essential internal biomarkers indicated and B) in vivo organization of the primitive heart tube and adult heart demonstrating the maturity of cells at early stages of differentiation(*1*).

Cardiomyocytes were first produced in human embryonic stem cells hESCs in 2001 by the Gepstein group in Israel(*64*). Further research elucidated the essential factors propelling embryonic stem cells down the mesodermal lineage by treatment with BMP to increase the percentage of CMs in each batch(*65*). Immediately after the discovery of induced pluripotent stem cells (iPSCs), the state of the art protocol was applied to turn these cells into CMs(*66*). Since then, a variety of protocols incorporating 3D cues(*67*), precise temporal control of chemical stimuli (*62*), and co-cultures (*68*) have been developed to increase the number of CMs produced from hESCs. These efforts have significantly improved the number of CMs which can be produced as measured by flow cytometry or other destructive forms of measurement.

Several companies now offer CMs produced in commercial processes, usually identified through the insertion of a fluorescent protein into a cardiac gene such as myosin heavy chain. These cells have been used by various diagnostic companies to establish programs such as CardioCHECK(*69*) and for preclinical assays which many in the pharmaceutical industry now regard as more predictive than animal models(*70*). Cells believed to be CMs have also been used in several clinical trials, notably the ALCADIA, CADUCEUS, and SCIPIO trials (*57*). While CM cells are not currently believed to integrate into host tissues, there is a strong consensus that their paracrine effects are beneficial to the repair and regrowth of injured heart muscles(*71*).

Current challenges and opportunities in cardiomyocyte development

Cellular therapies have the potential to replace or regenerate damaged cardiac muscle, a novel therapeutic option that would help millions of patients a year in the US alone(*72*). Previous studies have shown that undifferentiated cells are unsuitable for

implantation due to the danger of teratoma formation(*73*), but several experiments and clinical trials have shown that the delivery of terminally differentiated cells is likely to be safe(*71*), and the next round of studies to demonstrate efficacy is now underway. While methodologies for generating cardiomyocytes (CMs) from hPSCs are rapidly improving(*62*), numerous hurdles stand in the way of their clinical use(*56*). All protocols which do not require genetic manipulation result in heterogeneous populations of cells which may not be as efficacious or safe as terminally differentiated cardiomyocytes (*58-60*). The current gold standard activinA/BMP-4 protocol yielded > 30% of CMs(*61*). Novel methods promise to deliver significantly higher yields of CMs with more than 40% of cells becoming CMs (Figure 7), however none of these techniques can match the purity of genetic manipulation techniques which can lead to up to 98% pure populations. One of the major challenges confronting the field is therefore the development of reproducible isolation techniques that allow scalable purification of cardiomyocytes. The considerable upscaling, effective enrichment, and purification methods should be developed for robust research and clinical use in the future.

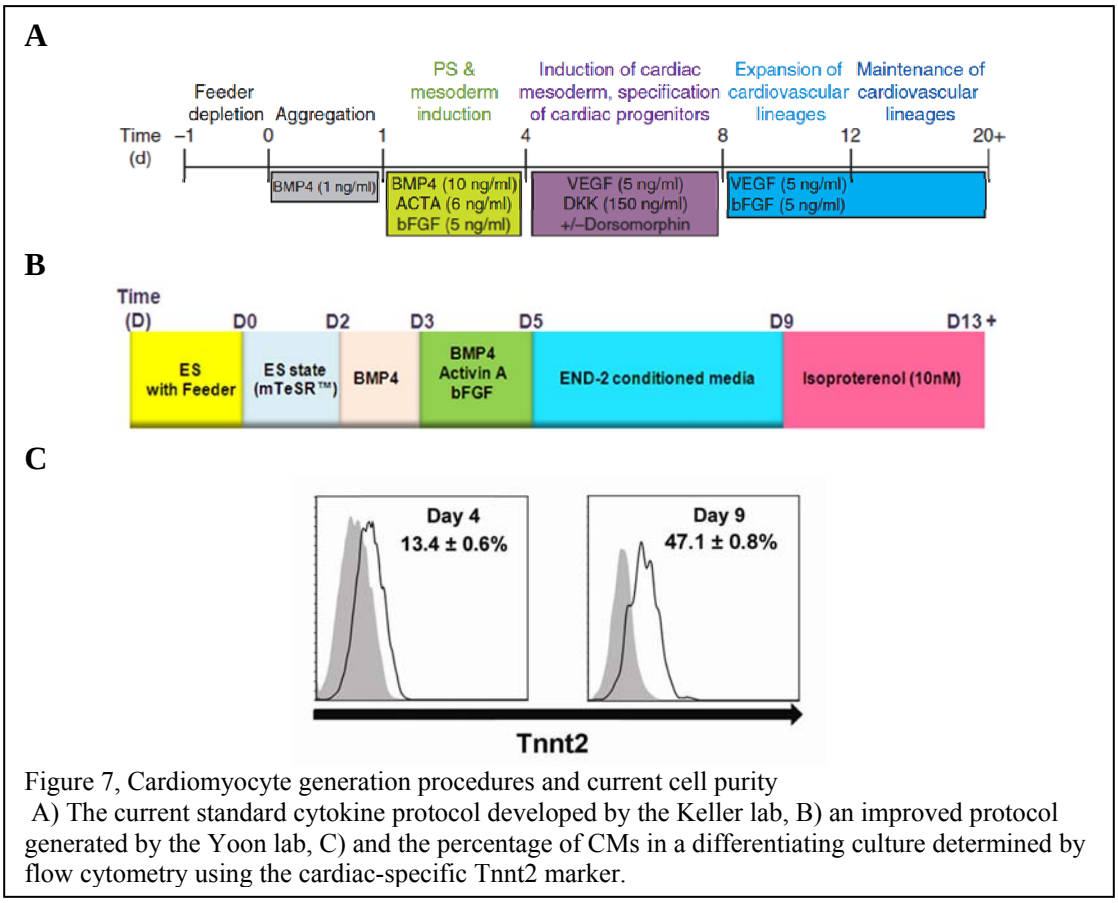

Figure 7, Cardiomyocyte generation procedures and current cell purity
A) The current standard cytokine protocol developed by the Keller lab, B) an improved protocol generated by the Yoon lab, C) and the percentage of CMs in a differentiating culture determined by flow cytometry using the cardiac-specific Tnnt2 marker.

A number of approaches have been developed for purifying CMs towards this end. The first approach is to use a fluorescent reporter gene driven by a CM--restricted promoter such as NKX2.5, ISL1 or MHC to isolate CMs as well as cardiac progenitor cells(*74*). Although this approach is efficient, it requires genetic modification and thus is not compatible with clinical use. The second approach is to isolate CMs by physical separation using a Percoll gradient(*75*). The purity of the cells is problematic due to the lack of biochemical markers involved in this method. The third approach is based on

selecting cardiac progenitors, but not mature CMs, using surface markers such as KDR(*76*) or PDGFRA(*62*). This population also contains endothelial and smooth muscle populations with a mixture of CMs. Another non-genetic method for isolating hPSC-derived CMs is based on the use of the mitochondrial dye tetramethylrhodamine methyl ester perchlorate (TMRM). Because this dye only functions in CMs with high mitochondrial density, it does not detect most of the immature CMs in the cultures. Lastly, the recent identification of surface markers expressed on human pluripotent stem cell (PSC)-derived CMs, EMILIN(*60, 62*), SIRPA(*77, 78*)and VCAM-1(*78*) allowed isolation of highly enriched populations of CMs by fluorescent activated cell sorting (FACS) or magnetic bead-based sorting (MACS®). However, these markers are not specific for CMs and their expression levels in the brain and the lung are higher than the heart, raising significant concerns regarding their authenticity as a single isolation marker for CMs(*78, 79*).

A method that allows tracking cardiac subtypes in a high throughput manner will accelerate research into the regulation of cardiac subtype specification. Currently, cardiac subtypes are distinguished by electrophysiology, a labor intensive and time consuming process that can only be performed by highly skilled electrophysiologists. Here we report the development of molecular beacon based methods for high throughput separation of cardiomyocytes and cardiomyocyte subtypes. Results from this study will not only significantly advance the application of hPSCs in safe cell therapy but also facilitate other applications of these cardiomyocytes such as in drug discovery and toxicity screening and in the study of heart development and diseases (*70, 80, 81*).

CHAPTER 3

DEVELOPING MOLECULAR BEACONS TO TARGET CARDIOMYOCYTE SPECIFIC MRNA

hPSCs show a huge potential for regeneration, however previous studies have shown that terminally differentiated cells must be implanted to avoid in vivo teratoma formation. The ALCADIO, SCIPIO, and CADUCEUS clinical trials have already produced Phase 1 results using terminally differentiated cells, which appear to indicate a reduction in fibrous scar tissue and without an accompanying increase in ventricular ejection fraction32. This may be due to the low percentage of cardiomyocytes contained in the implanted cells. A high throughput method for separating cardiomyocytes from differentiating cultures of hPSCs is therefore needed to implant high purity, terminally differentiated cardiomyocytes in clinical therapies.

A number of approaches have been developed for purifying CMs from differentiating cultures of hPSCs. The traditional approach is to use a fluorescent reporter gene driven by a CM--restricted promoter such as NKX2.5 (*78*), ISL1 (*82*) or MHC (*83*) to isolate CMs as well as cardiac progenitor cells. Although this approach is efficient, it requires genetic modification and thus is not compatible with clinical use. Another standard approach is to isolate CMs by physical separation using a Percoll gradient (*61, 75*). The purity of the cells is problematic due to the lack of biochemical markers involved in this method. A recently developed approach is based on selecting cardiac progenitors, but not mature CMs, using surface markers such as KDR11 or PDGFRA12. This population also contains endothelial and smooth muscle populations with a mixture of CMs. Another new method for isolating hPSC-derived CMs is based on the use of the

mitochondrial dye tetramethylrhodamine methyl ester perchlorate (TMRM) (*84*). Because this dye only functions in CMs with high mitochondrial density, it does not detect most of the immature CMs in the cultures (*77*). Lastly, the recent identification of surface markers expressed on human PSC-derived CMs, EMILIN2 (*85*), SIRPA (*77, 78*), and VCAM-1 (*78, 86*) allowed isolation of highly enriched populations of CMs by FACS or magnetic bead-based sorting (MACS®). However, these markers are not specific for CMs and their expression levels in the brain and the lung are higher than the heart, raising significant concerns regarding their authenticity as a single isolation marker for CMs (*78, 79*). The objective of this specific aim is to address these issues by developing a novel method to purify hPSC-derived CMs via MBs targeting mRNA of CM-specific genes uniquely expressed in CMs.

mRNA target selection and MB design

To determine optimal candidate genes detectable by MBs, we performed quantitative reverse transcriptase PCR (qRT-PCR) analysis on known cardiac specific genes and transcription factors using mRNAs extracted from freshly isolated mouse adult CMs and human neonatal heart tissues (Figure 8). Total RNA was prepared with the RNeasy mini plus kit (Qiagen) according to the manufacturer's instructions. The extracted RNA (100 ng to 1 mg) was reverse transcribed into cDNA (reverse transcription) via Taqman reverse transcription reagents including random hexamers, oligo (dT), and MultiScribe™ MuLV reverse transcriptase (Applied Biosystems). qPCR was performed on a 7500 Fast Real Time PCR system (Applied Biosystems) using Fast SYBR Green master mix (Applied Biosystems). All annealing steps were carried out at 60 °C. Relative mRNA expression of target genes was calculated with the comparative

threshold cycle (CT) method. All target genes were normalized to GAPDH in multiplexed reactions performed in triplicate. Differences in CT values (ΔCT = CT gene of interest−CT GAPDH in experimental samples) were calculated for each target mRNA by subtracting the mean value of GAPDH (relative expression = 2-ΔCT). The well

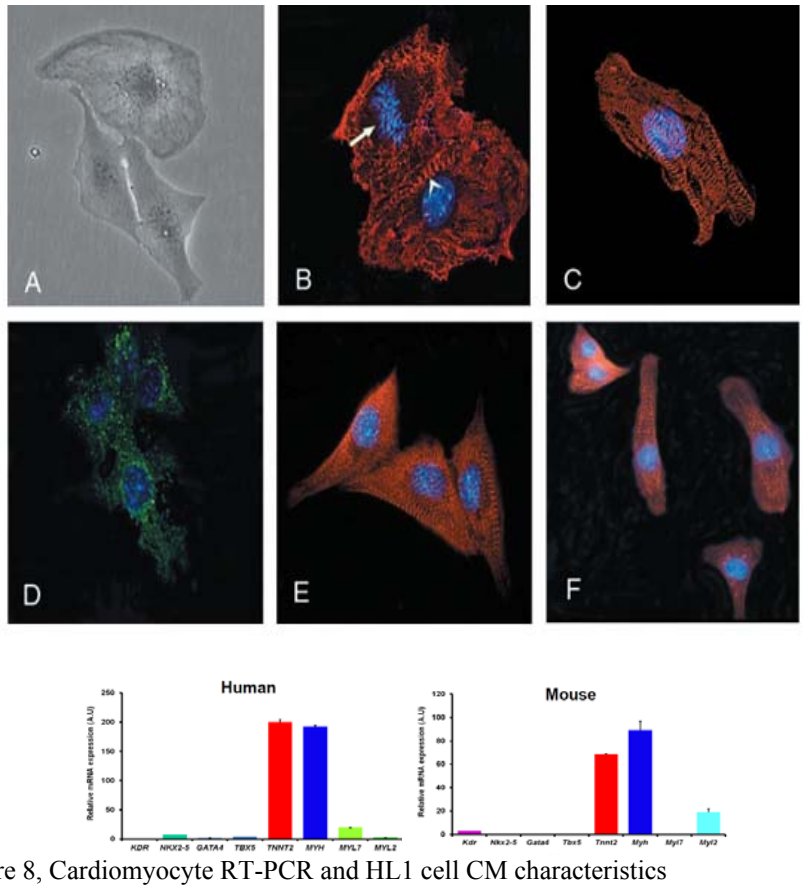

Figure 8, Cardiomyocyte RT-PCR and HL1 cell CM characteristics Morphological characteristics of HL1 cells with ICC staining demonstrating the presence of B) sarcomeric myosin, C) organized sarcomeric myosin, D) natriuretic peptide A, E) titin, F) and alpha-actinin. qRT-PCR using mouse adult CMs and human neonatal heart tissues.

known cardiac structural genes, cardiac troponin T (TNNT2, also known as cTNT) and myosin heavy chain (MYH6/7, also known as MHC) were most highly expressed in both samples and thus were determined as targets for MBs. We then decided to use the HL1a cell line (a generous gift from Dr. Claycomb) as a rapid in vitro testing platform. As shown in Figure 8 A-F, HL1a cells express several well known cardiac proteins and so provide a simple positive control system for beacon testing.

We designed five MBs (Appendix A) targeting unique sites in TNNT2 or MYH6/7 mRNA for both mouse and human cells using design rules determined by previous publications 38,39 and BLAST searches to ensure uniqueness. We predicted the secondary structure of our target mRNA molecules using mFOLD and identified areas which are unlikely to be double stranded in the cytoplasm. We further limited our search to areas of the mRNAs that were exactly the same in mouse and human genes. With all of these design constraints we arrived at a small number of molecular beacon target sites. These MBs were synthesized with a Cy3 fluorophore on the 5' end and a Black Hole Quencher 2 on the 3' end.

Verification of MB specificity

We quantified beacon fluorescence signal when hybridized to perfectly complementary and mismatched targets by incubating 500 nM molecular beacon solutions with target solution of increasing concentrations (Figure 9). Beacon signal was recorded using a microplate reader and normalized by the signal in wells with beacon only. All beacons displayed a linear response to increasing concentrations of complementary target and a low response to mismatched targets.

In order to determine the most efficient transfection method to deliver MBs into living cells we compared Streptolysin O (*38*), Lipofectamine 2000 (*87, 88*),

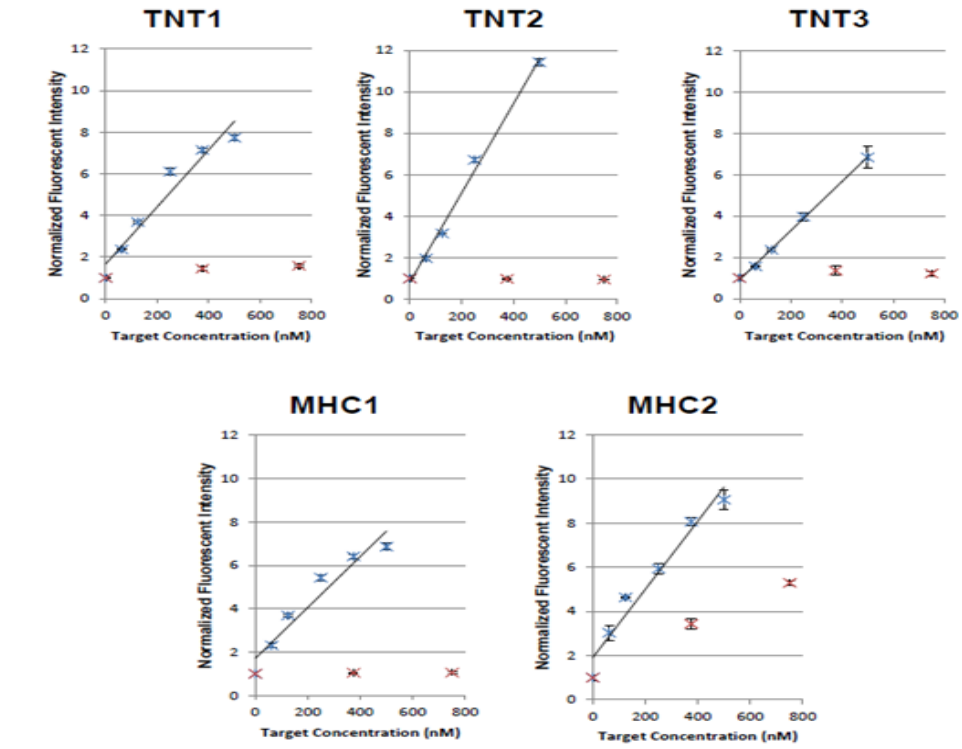

Figure 9, Verification of MB specificity.
Microplate data showing that MBs bind to complementary synthetic oligonucleotides but not those with mismatches in the sequence

Lullaby (*89, 90*), nucleofection (*91-93*), electroporation (*94*) and microinjection (*31*). To quantify the delivery efficiency, we designed an MB which has a nonspecific

sequence with FAM dyes conjugated to both 5' and 3' ends, allowing the probe to fluoresce regardless of its open or closed conformation. For all nucleofection studies, target cells were dissociated by treatment with Accutase (e-bioscience) and filtered through a 40-μm cell strainer (BD science) immediately before nucleofection. The dissociated cells (0.5-1 × 106) were carefully suspended in 100 μl of nucleofector solution V (Lonza) maintained at room temperature, and 0.5 μl of 500 nM MB was added for each reaction. Nucleofection was performed using a Nucleofector II (Amaxa Biosystems) set to the A033 nucleofector program. After termination of nucleofection, 500 μl of cold DMEM/F12 media was added to the reaction cuvette and the contents were gently transferred into a clean tube by a flexible pipette (Lonza). All procedures for nucleofection were performed inside a biological safety cabinet (Labconco) in the dark to prevent light induced nonspecific reaction of MB. Subsequently, 1 ml of pre-warmed DMEM/F12 media was added into each tube and further incubated in a 5% CO2

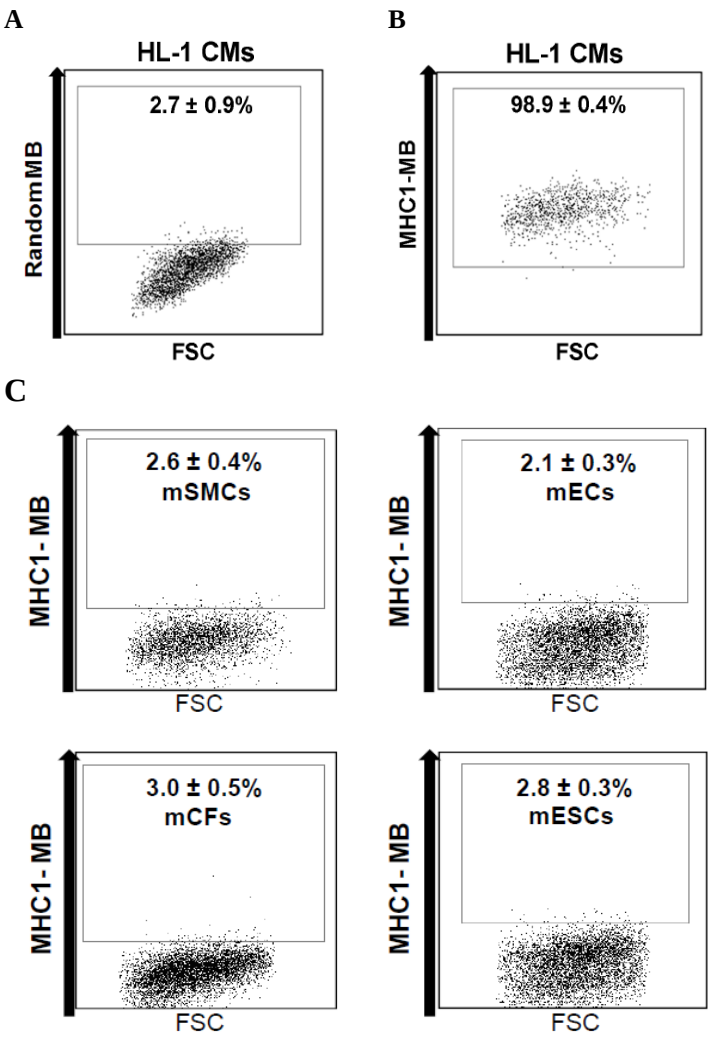

Figure 10, CMs cause an increase only in CM-specific MB fluorescence.
A) A random beacon delivered to HL-1 CMs does not cause an increase in fluorescence, however B) MHC-1 MB significantly increases its fluorescent intensity in HL-1 CMs. C) Flow cytometry showing that non-CM cell lines do not cause an increase in beacon fluorescent signal.

atmosphere at 37°C for 10 min for MB reaction. We delivered the nonspecific MB into a range of different cell lines including HL-1 CMs, smooth muscle cells, mouse embryonic fibroblasts and mouse embryonic stem cells. Flow cytometry analysis demonstrated that regardless of cell type, nucleofection consistently resulted in

internalized MBs in >95% of the cells, demonstrating the highest delivery efficiency among the methods tested. All FACS recordings were taken by centrifuging cells at 2000 rpm for 2 min, resuspended in DMEM/F12 basal media, and maintained on ice for 20 min to recover. Cells were then analyzed using a C6 Flow Cytometer (BD Biosciences) or sorted using a BD FACS Aria II cell sorter (BD Biosciences). Beacon signal was recorded using a 561 nm laser with a 585/15 nm emission filter to optimally excite and detect Cy3. Data were analyzed using FlowJo software (Treestar).

We determined the sensitivity and specificity of MBs designed to target CM-specific mRNAs using our HL1 test system. Each of the candidate MBs targeting TNNT2 or MHC mRNA was delivered into live HL-1 CMs by nucleofection and the efficacy was analyzed by flow cytometry. The results showed that one MB, designated as MHC1-MB, produced a much higher rate of MB signal positive cells in flow cytometry analysis (98.9%) compared to other MBs (TNT1, -2, -3 and MHC2). Microscopic fluorescence imaging also confirmed these results (data not shown). To determine the specificity of the MHC1-MB, a 'random'-sequence MB ('random beacon') which has a 16-base target sequence that does not match with any sequence in the mouse or human genome, was delivered as a negative control and displayed negligible fluorescence in HL-1 CMs (Figure 10). This ruled out the possibility that the fluorescence signal from MHC1-MB was due to nonspecific interactions and/or probe degradation by endonucleases. To further verify the specificity of the MHC1-MB, we delivered MHC1-MB into SMCs, mouse aortic endothelial cells (MECs), mouse cardiac fibroblasts (CFs) and mESCs, which are the most likely potential contaminating cell types in differentiating PSC cultures. Flow cytometry analysis showed that less than 5% of these cells displayed a

detectable fluorescence signal (Figure 10C). These results suggest high specificity of the MHC1-MB for detecting the CM lineage.

Characterization of beacon positive cells

To determine whether the MHC1-MB can be used to isolate CMs from differentiating mouse PSCs, our collaborators in Dr. Yoon's lab first established a system to efficiently differentiate mESCs into CMs. Flow cytometry analysis demonstrated that the percentage of Tnnt2-positive cells were 13.4% and 47.1% at days 4 and 9, respectively. Immunostaining further demonstrated that cells dissociated from beating clumps displayed CM-specific proteins such as Tnnt2, Tnni3 and α-Actn1 (also known as α-sarcomeric actinin), confirming their CM nature. The results with mouse iPSCs were similar (data not shown). For immunostaining assays, cells were fixed with 4% PFA for 10 min at room temperature, washed twice with PBS, and permeabilized with 0.1% Triton X-100 in PBS for 10 min. Cells were then blocked with 1% BSA in PBS for 60 min at room temperature and incubated with anti-ACTN2 (Sigma; 1:100), mouse anti-TNNT2 (NeoMarkers; 1:100), or rabbit anti-cTnI (Abcam; 1:100) at 4°C overnight. The cells were washed three times with 1% Tween 20 in PBS and incubated with anti-mouse IgG– Alexa Fluor 594 (Invitrogen; 1:1000) or anti-rabbit IgG–Alexa Fluor 488 (Invotrogen; 1:1000) in PBS for 1 h at room temperature. DAPI was used for nuclear staining. The samples were visualized under a fluorescent microscope (Nikon) and a Zeiss LSM 510 Meta confocal laser scanning microscope and LSM 510 Image software (CLSM, Carl Zeiss).

After establishing the differentiation system, we attempted to isolate CMs from differentiating mESCs using MBs. The differentiating mESCs at day 9 were nucleofected to deliver MHC1-MB and subjected to FACS. The percentage of cells positive for fluorescence signal from MHC1-MB was 49.2 ± 4.8% (Figure 11A). There was high concordance between the rate of Tnnt2-positive cells (47.1%) and MHC1-MB positive cells (49.2%) analyzed by flow cytometry, supporting the specificity of MHC1-MB for detecting mESC-derived CMs. Most importantly, 98.4% of these FACS-sorted MHC1-

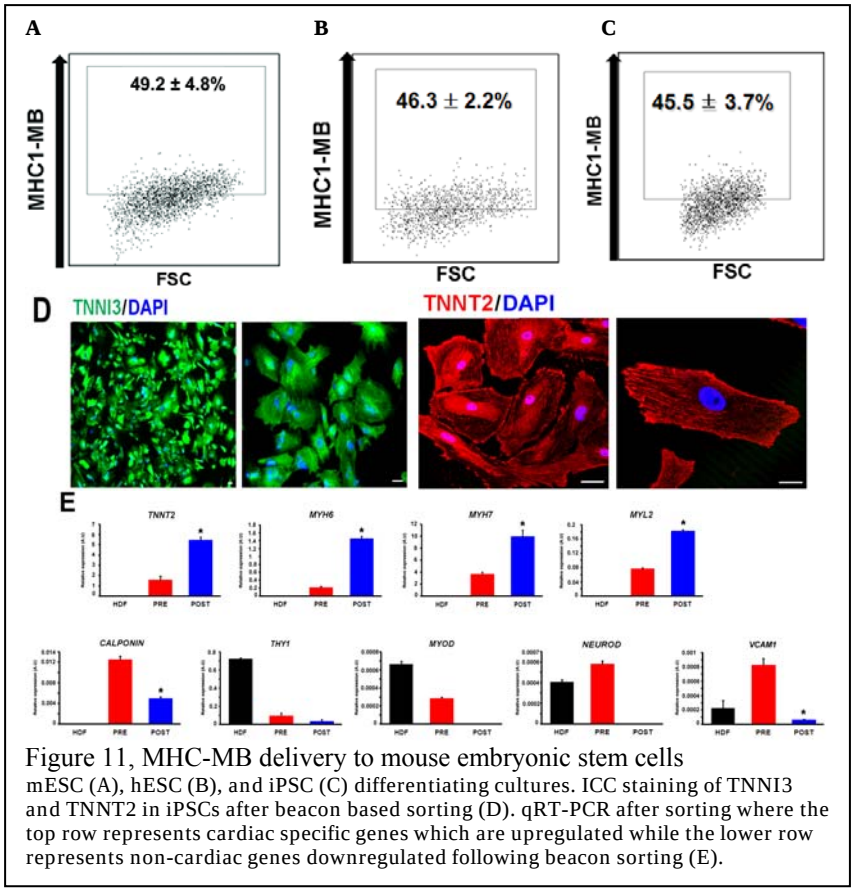

Figure 11, MHC-MB delivery to mouse embryonic stem cells
mESC (A), hESC (B), and iPSC (C) differentiating cultures. ICC staining of TNNI3 and TNNT2 in iPSCs after beacon based sorting (D). qRT-PCR after sorting where the top row represents cardiac specific genes which are upregulated while the lower row represents non-cardiac genes downregulated following beacon sorting (E).

MB positive cells exhibited Tnnt2 expression in flow cytometry. Immunocytochemistry further verified that virtually all MHC1-MB sorted cells stained positive for Tnnt2 and Actn2. qRT-PCR analyses demonstrated that these purified cells expressed at least 2 to 6-fold higher levels of Tnnt2, MHC, and Myl2 compared to the pre-sorted population. Other lineage genes were either expressed at negligible levels (Acta2, Ddr2, Gata4 and Sox17 or were non-detectable (Pecam1, Myogenic differentiation 1 (MYOD) and Neuro D) in the MB-purified CMs (data not shown).

Similarly, we investigated the utility of MB-based cell sorting for human PSC-derived CMs. Flow cytometry analysis demonstrated that the percentage of TNNI3-positive cells were 10.2% and 43.1% at days 9 and 13, respectively. We delivered MHC1-MBs to the cardiomyogenically differentiating hESCs at day 13. Flow cytometry analysis showed that the percentage of cells positive for MHC1-MB signal was 46.3% (Figure 11B), and these MHC1-MB positive cells exhibited TNNI3 expression at 97.6 ± 1.4% by flow cytometry. Almost all cells were stained positively for TNNT2 and TNNI3 and showed CM-like morphologies by immunocytochemistry. qRT-PCR analysis showed a significant increase in expression levels of CM-specific genes (TNNT2, MHC, and MYL2) and decrease in expression levels of genes specific for smooth muscle cells (SMC) (CALPONIN), fibroblast (THY1), skeletal myocyte (MYOD), neural lineages (NEUROD), and EC (PECAM1), suggesting enrichment of CMs and elimination of other lineage cells by cell sorting based on MHC1-MB signal.

We further determined the utility of MHC1-MB in isolating CMs from differentiating hiPSCs (BJ1; Gift from Dr. George Daley and Inhyun Park) (*95*, *96*). Our differentiation system for CMs from hESCs worked similarly to hiPSCs, yielding ~40.7%

of cTNI-positive CMs at day 13. MHC1-MB was delivered into differentiating hiPSCs at day 13 and the percentage of cells positive for MHC1-MB analyzed by flow cytometry was ~45.5%. Immunostaining with TNNI3 or ACTN2 identified almost all cells as CMs and qRT-PCR assay showed increased level of CM-specific genes and negligible levels of other lineage gene expression (Figure 11C).

Our results clearly show that MBs designed to target CM-specific mRNA in live cells can isolate functional CMs from differentiating mouse and human PSCs with high specificity and efficiency. We observed that the MBs targeting MHC identified up to 98% of an immortalized mouse CM cell line but less than 10% of the non-CM cells. When these MBs were delivered into both mouse and human PSCs derived CMs, 41 to 49% of the cells were identified as MB positive populations. More than 97% of the MB+ culture expressed cTNT/I determined by intracellular flow cytometry. Their identity as CMs was verified by immunocytochemistry and qRT-PCR, indicating that the beacon sorting did not disrupt processes essential to CM survival.

From a functional standpoint, stable action potentials (APs) were recorded from CMs purified via MHC1-MB that were cultured for 7-14 days after FACS sorting. Three major types of APs were observed such as nodal-like (6 of 46), atrial-like (11 of 46) and ventricular-like (29 of 46) APs (Figure 12a). These results indicate that cells purified via CM-specific MB are electrophysiologically intact and functional CMs and can maintain their functional characteristics in culture.

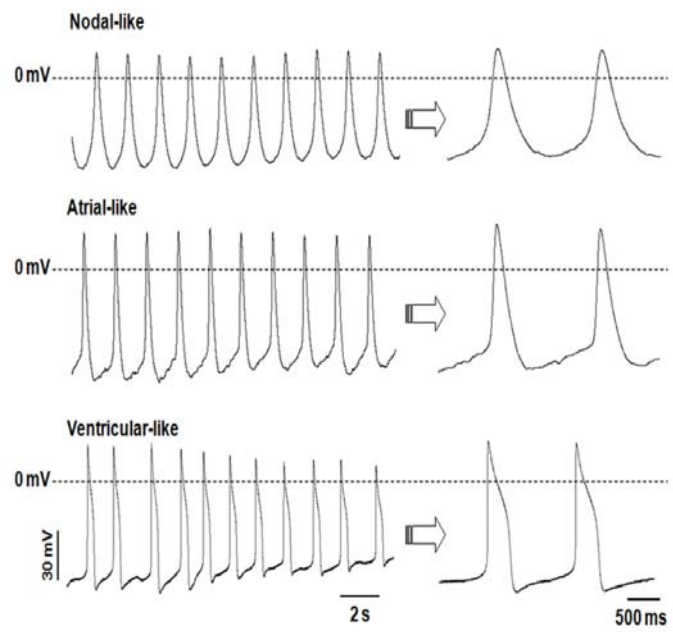

Figure 12, Functional characterization of sorted CMs
A) Representative electrophysiology recordings indicating the types of APs recorded from culture and B) *in vivo* implantation examination results clearly displaying tumorigenicity from unpurified cells.

To determine the behavior and effects of MB-based purified CMs in ischemic myocardium, purified or unpurified CMs derived from mESCs or the same volume of PBS were injected into the myocardium after induction of myocardial infarction (MI) in

mice. Echocardiography was performed weekly to measure cardiac remodeling and function. One week later, however, in the mice receiving unpurified CMs, a distinct mass was observed in the left ventricular lumen of the hearts, which grew over 4 weeks. Post-mortem examination at 3-4 weeks revealed tumor masses in 11 out of 12 mice (Figure 12B). By careful gross examination, tumors invaded internally into myocardium and externally into the pericardium. Cardiac tissues were fixed and stained with hematoxylin and eosin (H&E). Microscopic examination revealed that all tumors consisted of structures derived from all three embryonic germ layers, indicating teratomas. However, we did not detect tumors in any of the mice receiving MB-based purified CMs or PBS over the same follow-up period by echocardiographic or histologic examination. Tumor formation in unpurified-CM injected mice did not allow appropriate functional comparison between mice receiving unpurified- and purified CMs. However, purified CM injected mice showed a higher ejection fraction than PBS-injected mice, indicating improved cardiac function. We next conducted immunohistochemistry for cardiac tissues injected with purified CMs. Confocal microscopic examination demonstrated that injected CMs (DiI-positive) were engrafted as clusters, survived robustly for 4 weeks and expressed representative CM proteins. Taken together, these results suggest that injected MB-purified CMs are integrated into ischemic myocardium and are functional in vivo.

MB technology presents unique properties as a probe for enriching target cells based on mRNA expression levels. In contrast to genetic engineering approaches this technique can quickly and easily be applied to other highly expressed genes of interest and does not result in permanent alterations to the cell. This approach also has the advantage of detecting cardiomyocytes throughout the differentiation process, as MHC

and TNNT2 are both expressed continuously in adult cardiomyocytes. We predict that this technique will gain wide acceptance and usage in the broader academic community interested in generating high purity cardiomyocytes for a new generation of experiments to better characterize cardiac development.

CHAPTER 4

DEVELOPING MOLECULAR BEACONS TO TARGET VENTRICULAR CM-SPECIFIC MRNA

Heart failure is the leading cause of death worldwide, and current therapies including surgical and pharmacological interventions are only capable of delaying the progression of this detrimental disease(*97*). Particularly, patients suffering from MI which is one of the major causes of morbidity and mortality, have deteriorated ventricular heart function due to a loss of significant number of ventricular cardiomyocytes (CMs)(*98*). It has been well known that ventricular CM is the most extensively affected cardiac cell type when MI occurs(*99*). Due to the minimal ability of the adult mammalian heart to regenerate against lost or damaged ventricular CMs, there is great interest to identify potential cellular sources and strategies to regenerate new ventricular myocardium(*100*). Specifically, for treatment of MI, transplantation of a sufficient quantity of ventricular CMs, rather than other types of CMs such as nodal or atrial CMs, is preferred(*101*). Therefore, it is of great interest to generate a renewable source of ventricular CMs for cell-based therapies to treat MI patients.

Pluripotent stem cells including both embryonic stem cells (ESCs) and recently identified induced pluripotent stem cells (iPSCs) can self-renew and pluripotent, indicating they have the ability to develop into any type of cell, including CMs(*61, 102*). Importantly, several previous studies reported that PSC-derived CMs consist of distinct types of chamber specific CMs such as nodal, atrial, and ventricular(*61*). Hence, using *in vitro* differentiation methods, ventricular CMs generated from ESCs are considered as a

valuable source for potential cell-based cardiac regeneration therapy and as experimental models for human heart diseases. However, a problem is that there is no available method to selectively isolate ventricular CMs derived from PSCs, which typically consist of heterogeneous populations: atrial-, ventricular-, or nodal-like phenotypes as determined by electrophysiological analysis of APs. Since, the CMs comprising each heart chamber have unique functional, structural, and electrophysiological characteristics(*103*), it is obvious that transplantation with a mixed population of chamber specific CMs may cause arrhythmia and/or defective cardiac function due to inappropriate electromechanical integration of the grafted cells with the host myocardium(*104*).

Ventricular cardiomyocyte mRNA selection

Previously, a number of research groups have established transgenic murine ESC (mESC) and embryonic carcinoma cell lines(*105-107*), which have a fluorescent reporter gene driven by a Mlc2v (or Myl2) promoter, for the isolation of ventricular CMs. However, their approaches require genetic modification and therefore are incompatible for clinical use. Hence, there is high demand for the development of a system to selectively enrich a pure population of ventricular CMs, which is ideal for the purpose of cardiac regenerative therapy, as ventricular CMs are mainly responsible for cardiac contractile function.

Accordingly, we have developed a novel strategy to enrich ventricular cardiomyocytes (VCMs) from differentiating PSCs by targeting Iroquois homeobox gene 4 (Irx4), a transcription factor that is specifically expressed in ventricular tissue but not in atrial tissue(*108, 109*). We hypothesized that MBs hybridized to Irx4 mRNAs could enable isolation of VCMs from a mixed population.

Therefore, in the second aim of this thesis we first developed protocols to differentiate mouse ESCs into CMs which included a substantial percentage of ventricular CMs and also devised a strategy to isolate these ventricular CMs by applying MBs targeting ventricular CM-specific mRNAs followed by FACS sorting. Our results suggest the possibility of purifying ventricular CMs at a high efficiency and specificity from PSC differentiation cultures with this innovative and clinically compatible purification system.

We selected Irx4 after an extensive literature search as a target gene for generating ventricular CM-specific MBs (*108-110*). We then measured mRNA expression levels of Irx4 via quantitative RT-PCR (qRT-PCR) analysis in CMs isolated from either ventricles or atria of mouse adult hearts. We also measured Myl2, which is a well-defined ventricular CM-specific gene, as a positive control. The results showed that Irx4 is robustly expressed in ventricular CMs but not atrial CMs. The expression levels of both Irx4 and Myl2 mRNAs were substantially higher in mouse ventricular CMs compared to atrial CMs, indicating that Irx4 is a viable target for MB selection.

Molecular beacon development for the detection of ventricular CM mRNA

We designed four IRX4 MBs targeting distinct sites on the mouse Irx4 mRNA using design rules optimized in our previous publications 24, 26, 28. In addition, we used mFold (*111*) and the RNA Composer Webserver (*112*) to model our IRX4 MBs and to predict binding site availability in the target mRNAs (Figure 13A). These IRX4 MBs were synthesized with a Cy3 fluorophore on the 5’ end and a Black Hole Quencher 2 on the 3’ end as specified in Table 1. We quantified MB fluorescence signals when hybridized to perfectly complementary or mismatched synthetic targets by incubating 500

nM MB solution with targets of increasing concentrations (60-500 nM). IRX4 MB signals were recorded using a microplate reader and normalized to the signals in wells with beacon only. All IRX4 MBs displayed a linear response to increasing concentrations of complementary targets and a low response to mismatched targets (Figure 13B).

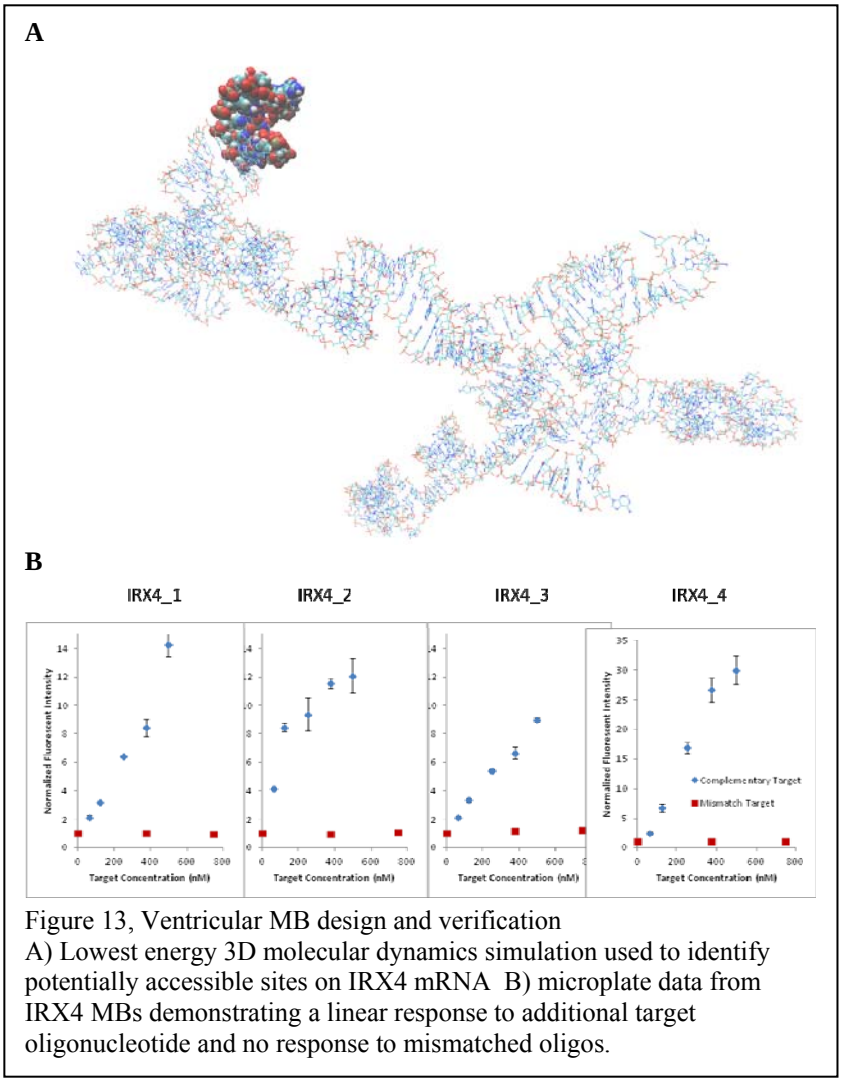

Figure 13, Ventricular MB design and verification
A) Lowest energy 3D molecular dynamics simulation used to identify potentially accessible sites on IRX4 mRNA B) microplate data from IRX4 MBs demonstrating a linear response to additional target oligonucleotide and no response to mismatched oligos.

Delivery of MBs into the cells

After attempting several different delivery methods for transfection of MBs, we found that nucleofection with nucleofection solution V and program A033 (Lonza) was an efficient method to deliver MBs to a variety of cell types28. To further refine this system, we also designed two distinct MBs to act as controls. A nonspecific interaction indicator MB (RQ) contains a random 20 bp loop sequence which is not similar to any known RNA in mouse cells, so any fluorescence would indicate an undesirable mechanism. A confirmation of delivery control MB (UQ) used the same random sequence but does not contain a quencher so that it fluoresces at all times. Both MBs were delivered to cells to ensure that transfection was efficient and that MB signal was specific to the targeted sequence.

Optimal Beacon Selection

In order to select the best MB to efficiently identify ventricular CMs, we examined the specificity, sensitivity and reliability of each IRX4 MB in three separate systems. To determine the specificity of MBs, we delivered each of the IRX4 MBs into mouse embryonic fibroblasts (mEFs), which do not express IRX4, by nucleofection, and analyzed the cells which showed false positive signals using flow cytometry. Among the three MBs (IRX4-1, -2, and -3) examined, the MBs designated as IRX4-1 and IRX4-2 yielded significantly lower rates of MB positive cells on mEFs (IRX4-1: 0.9 ± 0.1%, and IRX4-2: 1.4 ± 0.1%). In contrast, IRX4-3 displayed a high percentage of false positive signals (IRX4-3: 40.6 ± 3.8%) in mEFs. Hence both IRX4-1 and IRX4-2 were selected for further experiments and IRX4-3 was eliminated from the list of candidates

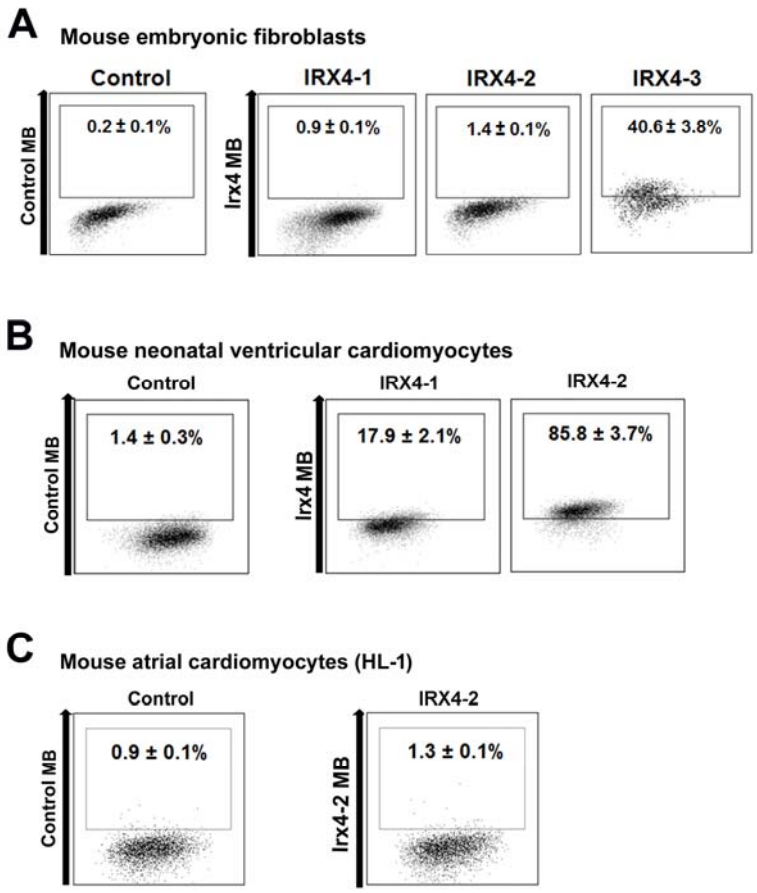

Figure 14, Optimal ventricular MB selection
A) Ventricular MB selection in embryonic fibroblasts after delivering three different IRX4 MBs designed to identify Irx4 mRNAs into mouse embryonic fibroblasts. N = 3. (B) Flow cytometric analysis of neonatal mouse ventricular CMs treated with control MB versus two IRX4 MBs targeting Irx4 mRNAs. N = 3. (C) Flow cytometric analysis of HL-1 CMs treated with control MB versus IRX4-2-MB. N = 3. The numbers in each panel represent the percentages of fluorescent cells.

(Figure 14A). Next, to evaluate the sensitivity of the two candidate MBs (IRX4-1, and -2), each of the candidate IRX4 MBs was delivered into mouse neonatal ventricular CMs and analyzed with flow cytometry. IRX4-2 MB identified a substantially higher percentage of ventricular CMs (85.8 ± 3.7%) compared to IRX4-1 MB (17.9 ± 2.1%).

Based on these results, we decided to use IRX4-2 MB as a final candidate for enriching the mESC-derived ventricular CMs (Figure 14B).

To further test its specificity, we examined the IRX4-2 MB with HL-1 CMs, an immortalized mouse atrial CM cell line known to retain atrial CM characteristics31, 32. The results from flow cytometry analysis demonstrated that less than 2% of HL-1 CMs displayed positive signals for IRX4-2 MBs, further supporting their specificity for ventricular CMs (Figure 14C). We next tested the IRX4-2 MBs against mouse SMCs, mouse aortic endothelial cells (mECs), mouse cardiac fibroblasts (mCFs) and mESCs, which are the most likely contaminating cell types in cardiomyogenically differentiated PSC cultures. Flow cytometry analysis showed that fewer than 3% of those cells displayed detectable fluorescence signals. These results clearly demonstrate that IRX4-2 MB is specific for identifying ventricular CMs.

Generation of ventricular CMs from mouse ESCs

To ensure stable production of mESC-derived ventricular CMs, we first established an embryoid body (EB)-mediated CM different system. Briefly, undifferentiated mouse ESCs (J1) maintained on mouse embryonic fibroblast (STO) feeder cells were enzymatically detached to form EBs. Since EB-induced differentiation alone is not sufficient for producing a high percentage of CMs, we plated day-4 EBs into a fibronectin coated dish and added ascorbic acid (50 μg/ml) to enhance CM differentiation. Spontaneously beating clumps began to appear 3-4 days after plating. After 7 days of CM differentiation on monolayer cultures we enzymatically dissociated the cells and applied them to a discontinuous Percoll gradient (40.5% to 58.5%) to enrich mESC-derived CMs. Percoll mediated-separation typically produces three layers of cells

and the bottom layer was reported to include a higher percentage of CMs34. Thus, the cells in the bottom layer were collected and cultured for another 7 days in the presence of cyclosporine A (30 µg/ml) to further induce CM differentiation35.

qRT-PCR analyses revealed dynamic changes in the expression of CM-specific genes in our differentiation system indicative of efficient CM differentiation. Expression of cardiac contractile genes (Tnnt2 and Mhc), and genes for atrial (Mlc2a), and ventricular CMs (Myl2 and Irx4) began to appear 7 days after culture. Expression of Myl2 and Irx4 continuously increased until day 18 (Figure 15A).

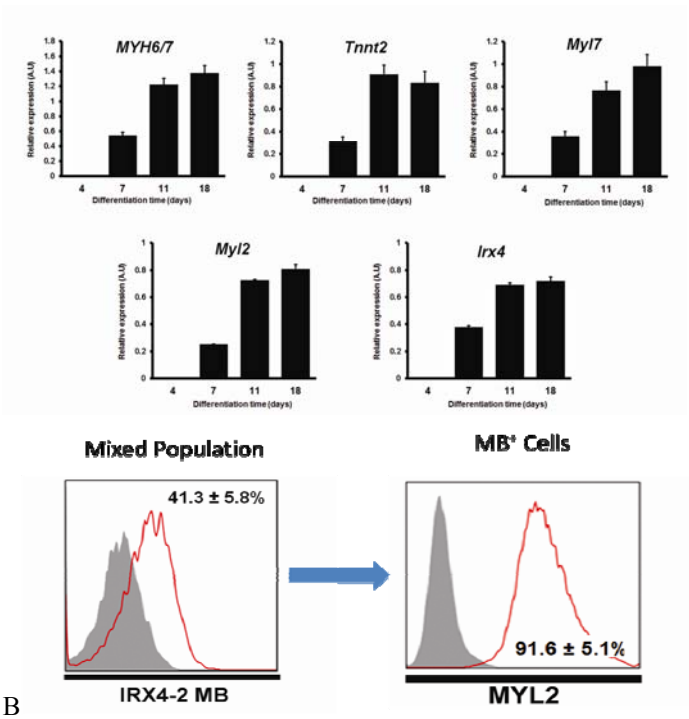

Figure 15, Differentiation of mESCs into ventricular CMs and impact of MB sorting

We next carried out immunocytochemistry and flow cytometry to further characterize the cell population at day 18. Immunocytochemistry demonstrated that day 18 cells significantly expressed CM-specific proteins including ACTN2 (α-sarcomeric actinin), TNNT2 (cardiac troponin T), and MHC6/7 confirming their CM nature. A substantial number of cells which were positive for ACTN2, TNNT2 and MYH6/7 concomitantly expressed MYL2 (or MLC2V) which is a specific protein for ventricular CMs. Flow cytometry analyses showed that the percentage of TNNT2 and MYL2 positive cells were 67.9 ± 4.5% and 39.2 ± 3.8%, respectively. These results clearly indicate efficient generation of CMs with a significant percentage of ventricular CMs through our CM differentiation system.

After establishing the CM differentiation system, we delivered IRX4-2 MB to the 18-day differentiated cells to isolate ventricular-like CMs. We used a pre-validated nucleofection protocol to deliver the MB and FACS-sort the cells. Flow cytometry results showed that 41.3 ± 5.8% of cells were positive for IRX4-2 MB fluorescence signal (Figure 15B). This number is similar to the detection rate of ventricular CMs using antibody-based (39.2% of Myl2-positive cells) methods. We then conducted FACS sorting for IRX4-2 MB and the MB positive CMs were seeded onto fibronectin coated plates for further experiments.

Characterization of beacon positive cells

The IRX4-2 MB positive CMs began to beat spontaneously within 48 hrs and continued to beat vigorously for up to two weeks. Only a small number of cells beat in the IRX4-2 MB-negative plate. 1-3 days after FACS sorting, we conducted flow cytometry analyses using TNNT2 and MYL2 antibodies to quantify the percentage of

CMs and ventricular-like CMs in IRX4-2-MB positive cells. Almost all sorted IRX4-2 MB positive cells expressed TNNT2 and MYL2 (TNNT2: 97.2% ± 3.4% and MYL2: 91.6 ± 5.1%) indicating efficient enrichment of mESC-derived ventricular CMs by IRX4-2-MB sorting (Figure 15B).

Electrophysiological characteristics of sorted cells

To investigate the electrophysiological characteristics of IRX4-2 MB-purified ventricular CMs, we performed whole-cell patch clamp analyses (Figure 16). It is known that electrophysiological characteristics of ventricular or atrial CMs are distinct. In primary

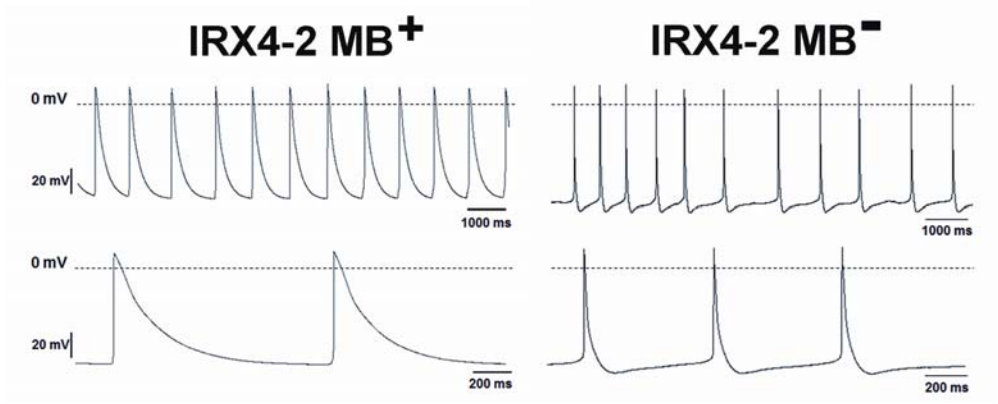

Figure 16, Electrophysiology of cells sorted based on IRX4 MB fluorescence
(A) Action potentials of the IRX4-2 MB-positive and -negative cells. Shown are representative configurations of the action potentials from IRX4-2 MB-positive (A**)** and negative cells (B). Action potentials were measured from 50 cells from either IRX4-2 MB-positive or -negative cells

adult CMs, average action potential duration (APD) of ventricular CMs is significantly longer than APDs of atrial CMs (*113*). The primary CMs in fetal stage showed similar patterns of APD at 90% of repolarization (APD90; ventricular CMs: 140 ± 7 ms vs. atrial CMs: 95 ± 7 ms) (*114*). In addition, CMs derived from mESCs displayed similar differences in electrophysiological characteristics between ventricular CM (APD50: 99.3

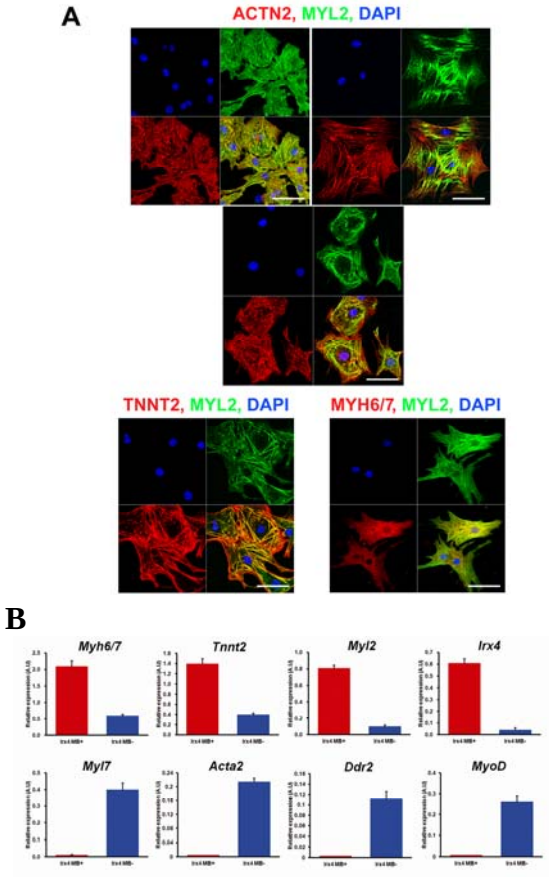

Figure 17, RNA and protein expression profiles of sorted and unsorted cells
(A) Immunocytochemistry for Actn2, Tnn2, and Myh6/7 on IRX4-2 MB-positive cells sorted from mESC cultures. Scale bars, 20 µm. **(B)** mRNA expression of cardiac and non-cardiac genes in IRX4-2 MB-positive and -negative cells measured by qRT-PCR. Y axis represents relative mRNA expression of target genes to GAPDH. *P < 0.05 compared to pre-sorted group. N = 3.

± 15.9 ms, APD90: 151.4 ± 22.1 ms) and atrial (APD50: 20.7 ± 7.2 ms, APD90: 60.6 ± 17 ms) (*115*). We observed that the APD50 of IRX4-2-MB positive and negative CMs were 159 ± 21.7 ms and 35 ± 7.8 ms, respectively ($P < 0.01$). Furthermore, the average upstroke slope of IRX4-2-MB positive and negative CMs were 251 ± 43.7 and 126.7 ± 51.9 mV/ms, respectively ($P < 0.01$) (Figure 16). Based on these results, we concluded that 98% of IRX4-2 MB positive cells possessed ventricular-type APs (49 out of 50 cells) whereas atrial- or nodal-type APs were not observed. Approximately 24% of IRX4-2 MB-negative cells exhibited an atrial-like AP (12 out of 50 cells).

To determine the contractile properties of IRX4-2 MB-based sorted CMs, we performed real time intracellular calcium [Ca2+]i imaging analysis. Spontaneous calcium transients were recorded where increasing calcium is measured and fluorescence intensity is normalized to the baseline measured at time 0 (Fo). Average calcium intensity along each point of the line indicates a cyclic calcium transient. Collectively, these results clearly show that the majority of the enriched IRX4-2 MB positive cells were functionally intact ventricular CMs.

Biochemical characterization of FACS-sorted ventricular CMs

To examine the cardiac identity and homogeneity of the ventricular CMs purified with IRX4-2 MB, immunocytochemistry was conducted with antibodies against various CM-specific markers (ACTN2, TNNT2 and MYH6/7 and MYL2) and ventricular CM marker (MYL2) 2-3 days after FACS sorting and cultures. As shown in Figure 17A, immunocytochemistry demonstrated that almost all IRX4-2-based enriched ventricular CMs exhibited ACTN2, TNNT2, and MYH6/7. Furthermore, a positive immunoreactivity for MYL2 was found in all of the IRX4-2 MB positive ventricular

CMs (Figure 17A). Importantly, these IRX4-2 MB positive ventricular CMs abundantly expressed GJA1 (known as connexin 43), an important connexin isoform in the formation of gap junctions between ventricular CMs, indicating that these enriched ventricular CMs possess the functional capability of forming cardiac junctions. qRT-PCR analyses further demonstrated that expression of ventricular CM genes Irx4 and Myl2 as well as general CM genes were substantially enriched in IRX4 MB positive cells compared to the IRX4 MB-negative cell population (Figure 17B). Furthermore, these IRX4 MB positive cells showed a significant increase in expression of general CM-specific genes (Tnnt2, and Myh6/7) compared to the IRX4 MB-negative cells (Figure 17B). Genes representing atrial specific CMs (Myl7) or other cell types were either expressed at negligible levels (Acta2, Ddr2, and MYOD) or were non-detectable (Pecam1, and Neuro D) in the IRX4 MB positive.

Taken together, the results clearly demonstrate that IRX4-2 MBs that target ventricular CM-specific mRNA in live cells allow isolation of functional ventricular CMs from differentiating mESCs with high specificity and efficiency. Despite the importance of ventricular CMs for cell-based therapy and drug development, no studies reported isolation of ventricular CMs from genetically unmodified stem cells. We generated homogeneous ventricular-type CMs from mESCs, without altering genomes, via MB-based sorting for Irx4, a ventricular CM-specific transcription factor. This method yielded functional ventricular CMs with high specificity and efficiency. Nucleofection of a selected MB targeting mRNA of Irx4 followed by FACS sorting enabled enrichment of ventricular-type CMs to 92% from differentiating mESCs. These purified CMs demonstrated ventricular CM-like APs at ~98% and Ca2+ oscillations in

electrophysiological studies, suggesting functionally intact ventricular CMs. These cells showed coordinated contraction and survived more than 2 weeks in culture while maintaining their phenotype.

Over the past decade, there has been notable advancement in the methodologies for generating PSCs (*116, 117*)and producing CMs from PSCs, raising our expectations for using stem cell-derived CMs for cardiac repair (*61, 102*). However, all reported CM differentiation protocols generate heterogeneous CMs mixed with other cell populations. Several recent studies reported non-genetic methods for isolating general CMs (*77, 118*). However, these protocols still generate heterogeneous CMs, not chamber specific CMs. Given the major role of ventricular CMs for cardiac contractile function, it would be important to develop a non-genetic method to isolate ventricular CMs from differentiating PSCs, which will be of value to many preclinical and clinical applications.

Studies have reported isolation of ventricular CMs derived from transgenic mESCs (*105, 107*) or embryonic carcinoma cell lines (*74*). For example, ventricular CMs were isolated from genetically modified mESCs in which a fluorescence reporter gene was driven by a promoter of Myl2 (*105, 107, 119, 120*). While useful for basic research, these genetic methods cannot be used for clinical applications or disease modeling due to genetic modifications. Cell sorting with specific surface markers is a preferred method for isolating target cells from PSC cultures. However, no unique surface markers are known for ventricular CMs, and to identify and validate such markers and develop antibodies would require considerable resources.

Our strategy was to directly target intracellular mRNAs of known ventricular CM-specific genes. This approach avoids genetic modifications and the need to identify

specific surface markers. We focused on Irx4, which is known to be a more specific identifier of ventricular CMs than Myl2 (*121, 122*). Irx4 plays a critical role in regulating chamber specific gene expression in the developing heart and is involved in determining ventricular cell specification (*108, 109*). Its expression is restricted to the ventricles throughout the developmental as well as postnatal periods (*108, 123*).

The uniqueness of this study is in being the first to target a transcription factor for isolating target cells. MB technology marks desired cells by targeting specific mRNA sequences (*124-126*). Recently we reported that this technology allowed sorting of general CMs from differentiating human and mouse PSCs (*127*). However, in that study, the targets were mRNAs of structural proteins such as MYH and TNNT2. As mRNAs of structural proteins are present in abundant copies, the likelihood of successful isolation was relatively high. In fact, we initially tried another transcription factor, NKX2.5, as a target in that study, but were not able to sufficiently label general CMs. In this study, our improved design technology enabled us to isolate ventricular CMs by targeting a transcription factor. To prove the specificity in this study, we carried out extensive assays to ensure that the MB signal was a precise indicator of ventricular CMs based on the hybridization of the probe to IRX4. We tested the MBs in solution with synthetic target ODNs that varied from the ideal sequences by six bp, the closest that a BLAST search through the mouse transcriptome allowed. However, a major challenge was whether the designed MBs could be hybridized to a sufficient quantity of target mRNA sequences to allow sorting of ventricular CMs. Among three designed MBs, one of them, IRX4-2 MB, was bound sufficiently to allow cell sorting. Another difficulty was to find a delivery method to deliver enough MBs to sufficiently label target mRNAs. We tested seven

transfection methods to deliver MBs into living cells, Streptolysin O, Lipofectamin 2000, Lullaby, electroporation, microinjection and nucleofection, and found that nucleofection in a specific buffer induced maximal target delectability with minimal cell toxicity (data not shown). Cytotoxic effects of MB themselves within the cells were reported negligible. We and others have found that MBs degrade within a few hours in the cells so that their effects on cell viability or cell identity are insignificant (*31, 124-127*). Even with repetitive transfections of MBs, we did not observe phenotypic or functional changes, as evidenced by unaffected spontaneous contraction and immunocytochemistry assays.

This unprecedented production of homogeneous and functional PSC-derived ventricular CMs using non-transgenic approach will yield new avenues for clinical and research applications. First, a pure population of ventricular CMs generated by this method offers a safer and effective option for cell therapy and tissue engineering. The mixed populations of PSC-derived CMs are more likely to cause abnormal electrical activity16 or less efficient contractile function50. From a research perspective, the MB-purified ventricular CMs represent a powerful in vitro tool for disease investigation and drug discovery. They could be used for better defined in vitro cardiac disease models for genetic or idiopathic cardiac diseases such as long QT syndrome (*128, 129*). They can also serve as an in vitro model to test chamber specific effects of cardiac drugs (*130*). These purified CMs will yield more accurate genetic and epigenetic information through high throughput sequencing techniques. We also expect that this MB-mediated cell sorting method can be applied for isolating other cardiac cells such as nodal cells or atrial CMs. By eliminating the need for expensive efforts to identify surface markers and

generating antibodies, this new technology can be further expanded to isolation of other cell types from PSCs such as neuronal cells or pancreatic β-cells.

CHAPTER 5

ISOLATION OF NODAL AND WORKING CARDIOMYOCYTES USING MOLECULAR BEACONS

hPSCs can be differentiated into working (atrial and ventricular)- and nodal-type cardiomyocytes, which are suitable for different applications: enriched working-type cells without the contamination of nodal-type cells for repairing injured ventricular myocardium, and enriched nodal-type cells for developing a biological pacemaker to treat arrhythmias caused by cell loss or dysfunction in pacemaker tissues (e.g. patients with congenital heart defects). However, none of the existing methods generate homogeneous cells for a specific subtype, and methods for the isolation and enrichment of each subtype have not been well developed. Research on cardiac subtype specification has been challenging due to the lack of an analysis tool that can be used in a high throughput platform to distinguish and enrich cardiac subtypes. This chapter details progress in tracking, enriching and characterizing cardiac subtypes using MBs. This technology will allow us to generate highly homogeneous nodal- or working-type cell populations and therefore facilitate future applications of hPSCs in cardiac cell therapy as well as in drug discovery and toxicity screening. It can also help accelerate studies to understand the regulation of cardiac subtype specification from hPSCs, which could shed light on heart development and diseases.

Cardiac subtypes are currently distinguished by measuring the action potential of single beating cells using the patch clamp technique, which is labor intensive and time consuming (*114*). A higher throughput method will therefore be an important tool for the advancement in basic research and practical applications of cardiac subtypes. To this end,

we designed MBs to target cardiac subtype-specific mRNAs, evaluated methods for efficient delivery of these MBs into living cells, and determined their specificity and sensitivity using both fluorescence imaging and flow cytometry assays with known cell types. We quantified the expression levels of cardiac subtype genes in model cells as well as differentiated cells derived from hPSCs using both MBs and RT-PCR, and determined their correlation.

Cardiac subtypes have distinct molecular and cellular features, such as mitochondrial content (*118*), proliferation rate (*131*) and specific gene expression (*77, 86*), which can be used to confirm specificity of the subtypes isolated based on MBs. In addition, each cardiac subtype has its unique electrophysiological phenotype (*114*), an important criterion for the confirmation of subtype specificity. We plan use the MBs to enrich putative nodal and working cardiomyocytes. evaluate molecular and cellular characteristics, and then confirm their functional properties by electrophysiology. Specific experiments include the detection of mitochondrial content, proliferation capacity, expression of gap junction channels and other subtype-associated genes, and the measurement of APs. In addition, we will evaluate if simultaneous detection of multiple markers using the MBs is required for accurate identification of cardiac subtypes.

Identification of target mRNAs

Differentiation of hPSCs into cardiomyocytes was induced using an established growth factor-guided method (*61, 132*). Briefly, cells were plated on a monolayer before the differentiation protocol began. Cells were then incubated with a growth media supplemented with 100 ng/ml activin A and 10 ng/ml BMP4. Cells were incubated in the medium for 14 days, at which point the cells displayed spontaneous contraction in

culture. In addition between 40 and 60% of cells expressed alpha-actinin, a cardiomyocyte biomarker, as assessed by flow cytometry.

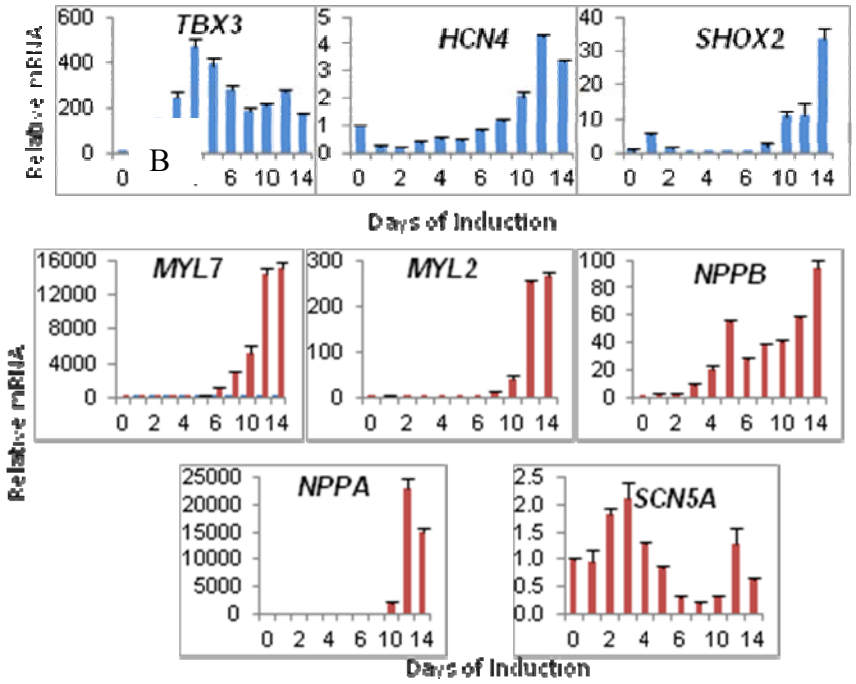

Figure 18, Expression of nodal and working genes in a differentiating culture A) Expression of nodal cell markers TBX3, HCN4, and SHOX2 in differentiating cells on each day of differentiation. B) Expression of working cell markers MYL7, MYL2, NPPB, NPPA, and SCN5A during the differentiation process.

RNA was isolated from cells each day during the differentiation procedure and 1 ug was converted to cDNA. The total cDNA was then analyzed using TaqMan RT-PCR

for the presence of nodal specific genes including SHOX2, TBX3, and HCN4 (Figure 18A) as well as working cell genes such as NPPA, NPPB, MYL2, and MYL7, and SCN5A (Figure 18B). As shown in Figure 18 the largest change in expression occurs in NPPA. The several thousand fold increase is likely to occur in approximately 40% of the cells, indicating hundreds, if not thousands of copies of this gene in each target cell.

In addition, several publications have shown that NPPA is ubiquitously expressed in immature working cells (*133*) but not expressed in the sunus venosus or atrioventricular canal of the primordial heart tube which develops into the cardiac conduction system (*1, 134, 135*). It is a secreted factor which is closely associated with myosin heavy chain and participates in the BMP4 cardiac differentiation pathway. Because of the abundant expression of NPPA in target cells and because of its central role in cardiac differentiation we decided to target NPPA for the isolation of working cells from a differentiating culture.

Molecular beacon development for the detection of working CM mRNA

We used a variety of MB designs to target NPPA. These include DNA MBs, Locked nucleic acid (LNA) MBs (*2*), and Reference Dye Molecular Beacons (RBMBs) (*3*). DNA MBs are the same as described in previous aims, however LNA MBs and RBMBs are alternative designs which offer potential advantages over the standard MB design. LNA is a variation on the standard RNA chemistry that significantly improves nuclease resistance and binding affinity due to the linkage of the 2' oxygen and 4' carbon of the sugar backbone (Figure 19A). MBs which incorporate LNAs have an increased affinity for their target sequence and a decreased tolerance for mismatches due to the high enthalpies involved in LNA Watson-Crick bonds. RBMBs were first developed in 2008

by Chen et al, and are advantageous due to the incorporation of an unquenched reference dye on each MB (Figure 19B). This allows each experiment to track the location and quantity of both quenched and unquenched beacons inside cells, and it allows the probe to use cellular RNAi transport mechanisms to be exported from the nucleus.

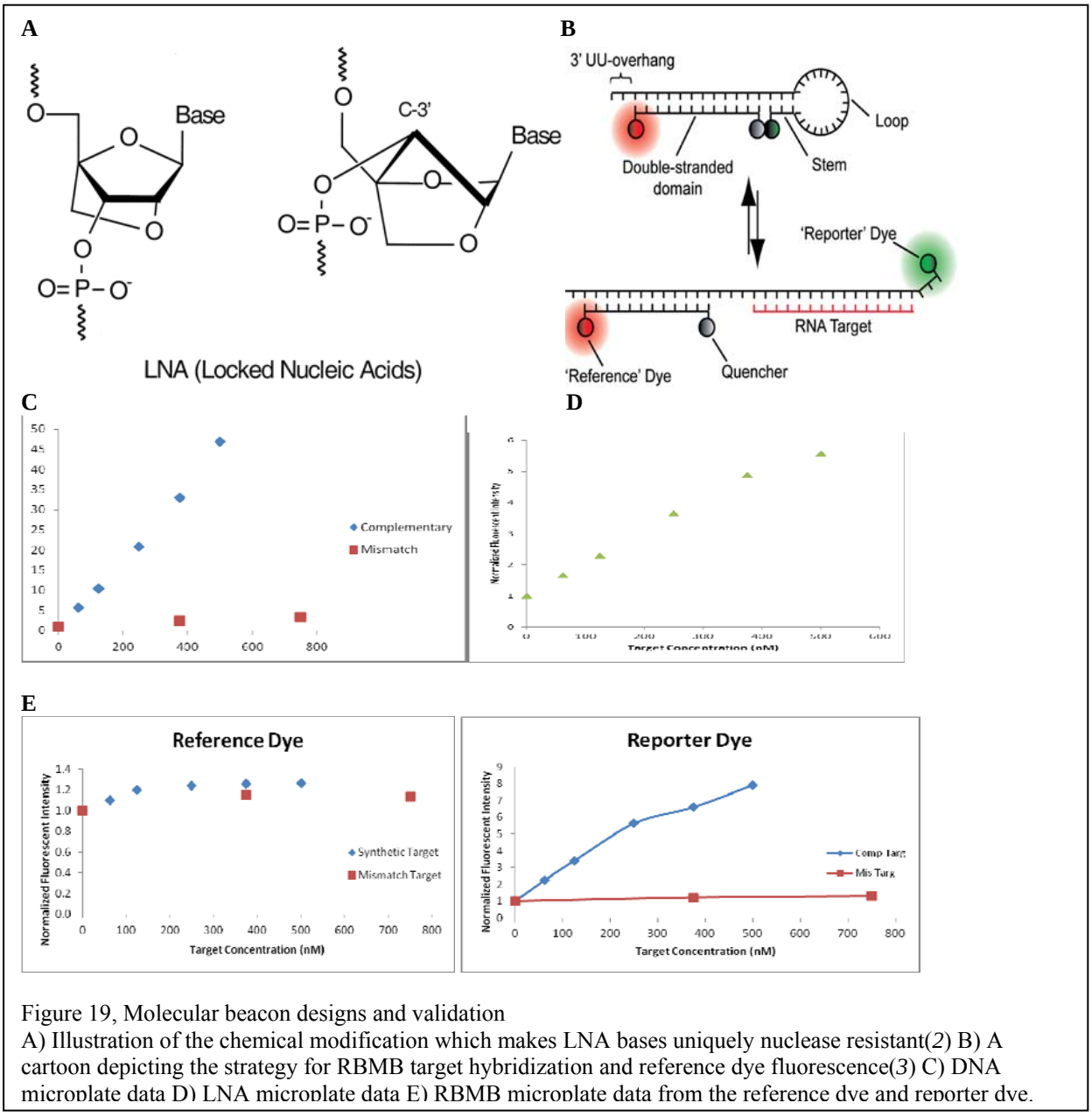

Figure 19, Molecular beacon designs and validation
A) Illustration of the chemical modification which makes LNA bases uniquely nuclease resistant(*2*) B) A cartoon depicting the strategy for RBMB target hybridization and reference dye fluorescence(*3*) C) DNA microplate data D) LNA microplate data E) RBMB microplate data from the reference dye and reporter dye.

mFold (*111*) and the RNA Composer Webserver (*112*) were used to model NPPA mRNA and to predict binding site availability in the target mRNAs. We designed 15 beacons targeting NPPA using DNA, LNA, and RBMB backbones (Appendix A). MBs were synthesized with a range of fluorophores and quenchers as described in Appendix A. We quantified MB fluorescence signals when hybridized to perfectly complementary or mismatched synthetic targets by incubating 500 nM MB solution with targets of increasing concentrations (60-500 nM). NPPA MB signals were recorded using a microplate reader and normalized to the signals in wells with beacon only. All NPPA MBs displayed a linear response to increasing concentrations of complementary targets and a low response to mismatched targets (Figure 19C-E). The signal to noise ratio for LNA MBs appeared to significantly increase in comparison to a similar DNA MB (Figure 19D), which matches predictions for an increased binding affinity to its target sequence. The reference dye in the RBMBs was also detected at equal intensities in each well (Figure 19E).

Development of positive control cell lines

HL1 cells would not be an ideal control line since all of these cells express atrial genes. In addition more precise positive control cell lines would differ only in the expression of the target gene. To recapitulate immature working and nodal cell conditions we decided to construct positive control cell lines by transfecting immortalized cell lines with plasmids expressing target genes. A plasmid expressing both NPPA and puromycin resistance genes was ordered from Origene and transformed into E. Coli to create a bacterial stock. The plasmid was maxi-prepped from the bacterial stock using a Qiagen Maxi-prep kit and then confirmed the sequence of plasmid through the custom

sequencing service from MWG. The plasmid was then transfected into CHO cells and 3T3 cells using Lipofectamine 2000. Briefly, 2 uL of Lipofectamine 2000 was incubated in 150 uL of Optimem for 5 minutes while 2 ug of plasmid was diluted into 150 uL of Optimem. The lipid and plasmid solutions were combined and allowed to complex for 20 minutes, after which the entire solution was added to a 6 well plate of cells (either 3T3 or CHO cells). Transfection was allowed to proceed for 24 hours before the media was replaced. A control plasmid expressing only GFP was also transfected to analyze transfection efficiency. 70-90% of CHO cells were GFP positive, while only 20-40% of 3T3 cells were GFP positive (data not shown) after the transfection procedure.

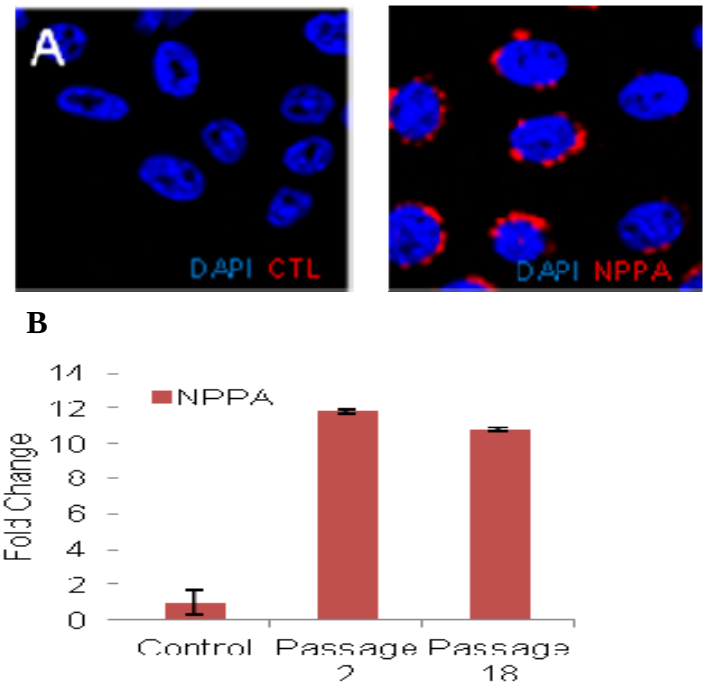

Figure 20, Characterization of NPPA positive control cell lines A) Immunocytochemistry data and B) RT PCR data demonstrating that stably transfected cells express significant quantities of NPPAduring extended cell culture

NPPA expressing cells were selected using 15 ug/ml puromycin for 1 week, during which time 100% of the GFP expressing cells died. Single cells were placed in each well of a 96 well plate and maintained under selective pressure for 3 weeks. RNA was isolated from surviving colonies and analyzed for NPPA expression. The colony expressing the most NPPA was selected and propagated as a positive control cell line. Cells were subjected to selective pressure (15 ug/mL puromycin) for 3 out of every 7 days to maintain stable gene expression.

The positive control cell line (CHO-NPPA) was analyzed using immunocytochemistry (Figure 20A), which indicated that 85-90% of cells in the culture expressed NPPA. In addition NPPA expression was analyzed at the second and 18^{th} passage after colony selection, showing a significant upregulation in comparison to untransfected cells and similar levels of expression. The cell line did not display significantly altered morphology or proliferation rates for at least 18 passages, indicating that the plasmid integration did not significantly disrupt essential cell maintenance pathways. Each cell should express several hundred copies of NPPA, providing a conservative predictor of the ability of MBs to discriminate between target mRNA and other mRNA in a cellular environment.

To test each NPPA beacon for sensitivity and specificity we transfected both NPPA-CHO and CHO cells with NPPA MBs. Briefly, 1 million cells were trypsinized, and the trypsin was neutralized with growth media containing 10% FBS. Cells were rinsed in PBS and then resuspended in nucleofection solution containing 500 nM MB. Cells were then nucleofected using the Amaxa Nucleofector program U024 and either plated onto gelatin coated glass slides or kept in suspension for flow cytometry. A

random quenched MB as described in Aim 2 was delivered to each cell type during each experiment to ensure that nonspecific effects were not causing MBs to open (Figure 21 A). In addition a random unquenched MB was delivered to assess delivery efficiency.

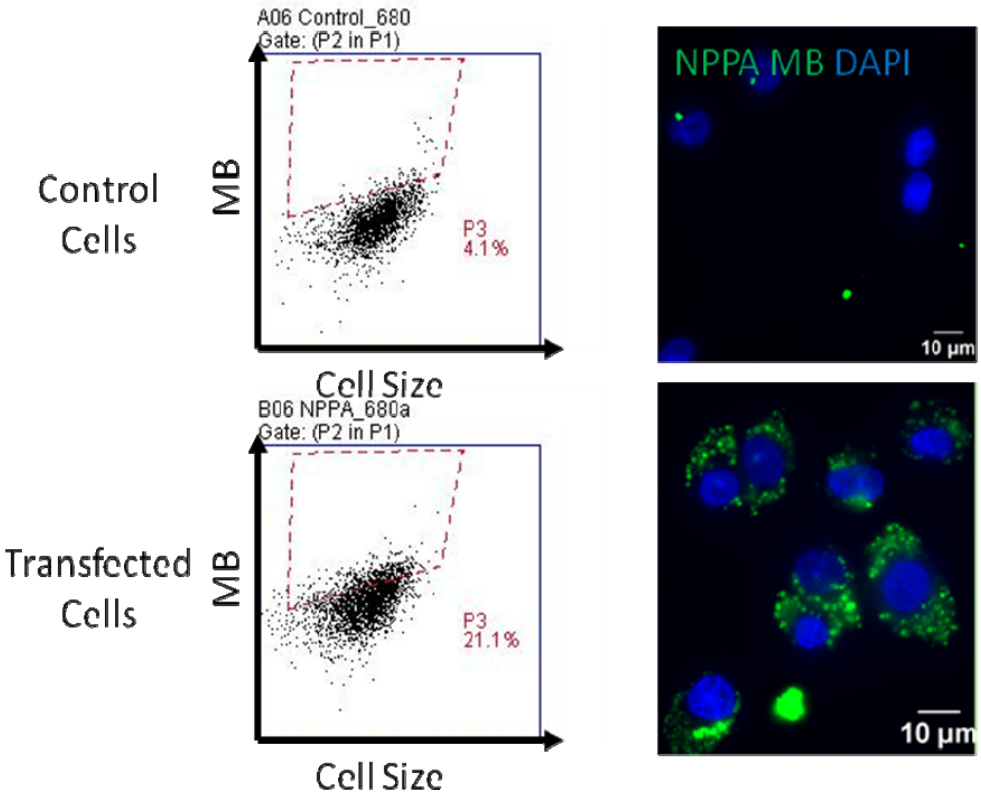

Figure 21, NPPA MBs in positive control cell lines
Flow cytometry data showing beacon signal in positive and negative control cells and fluorescence microscopy image of LNA NPPA MB in CHO-NPPA cytoplasm.

The NPPA 680 DNA and LNA beacons were observed to be more fluorescent in CHO-NPPA cells than in CHO cells (**Figure 21**). In addition fluorescence microscopy

was used to determine that the NPPA MB signal originated from cytoplasmic, rather than nuclear loci. NPPA MBs selected through this process carried a high degree of reliability because control beacons with identical ODN modifications behaved similarly in both CHO and CHO-NPPA cells, eliminating several systemic artifacts. In addition the only difference between CHO and CHO-NPPA cells was the expression of a single exogenous gene, which should be less drastic than the changes induced by differentiation systems used to test previous beacons. Finally, the detection of MB signal in cytoplasmic loci also reduced fears of systemic artifact which may be caused by nuclease activity, which occurs at the highest rate in the nucleus.

Delivery of MBs into stem cells

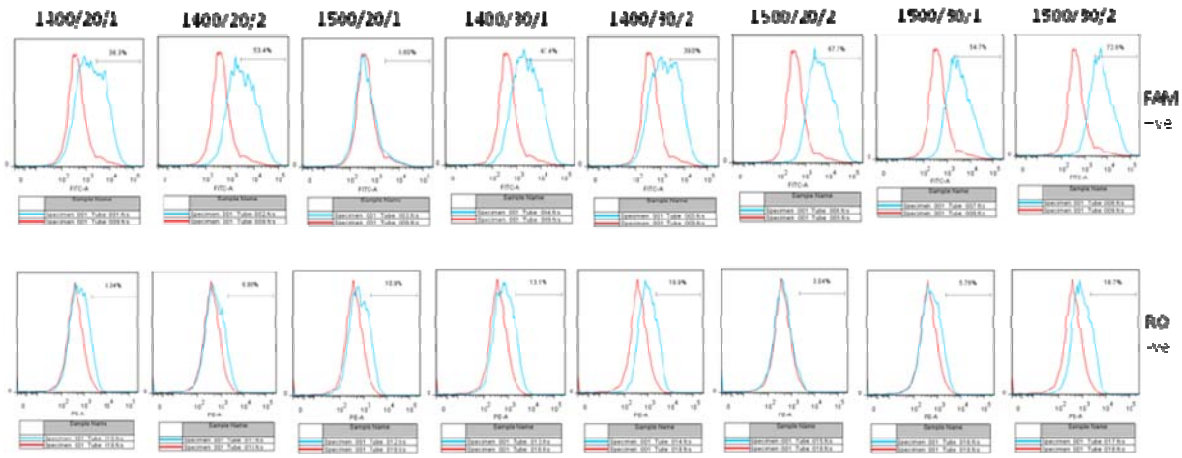

Figure 22, Optimization of Neon transfection in cardiomyocytes.
A) A wide range of transfection conditions were attempted, however we selected the 1500 V, 20 ms, 2 pulse conditions as well as the 1400 V, 20 ms, 2 pulse method due to their high transfection efficiency and low rate of false positives

Due to low cell yields in previous aims we decided to test out the Neon transfection system, recommended in several publications(*3, 28, 31, 33*) for stem cell transfections because of its reduction in cell death and high transfection efficiency. In addition the 10 uL Neon transfection kit allows the use of 100,000 cells per transfection, rather than the 1 million cells used for nucleofection. Due to the expense of stem cell culture and the long development time required for cardiomyocytes the order of magnitude reduction in cell numbers becomes an important consideration.

Towards this end we tested 30 different parameter sets based on optimization parameters suggested by the literature and the manufacturer. We used positive control MBs to evaluate transfection efficiencies using very similar types of probes, and false positives were assessed by transfecting random loop sequence quenched MBs. The random negative MBs served the dual purpose fo detecting increased cell size or increased nuclease activity which is associated with cell death or stress. We found two parameter sets, 1500/30/2 and 1400/20/2 which resulted in a high transfection rate, low false positive rate, and high cell viability.

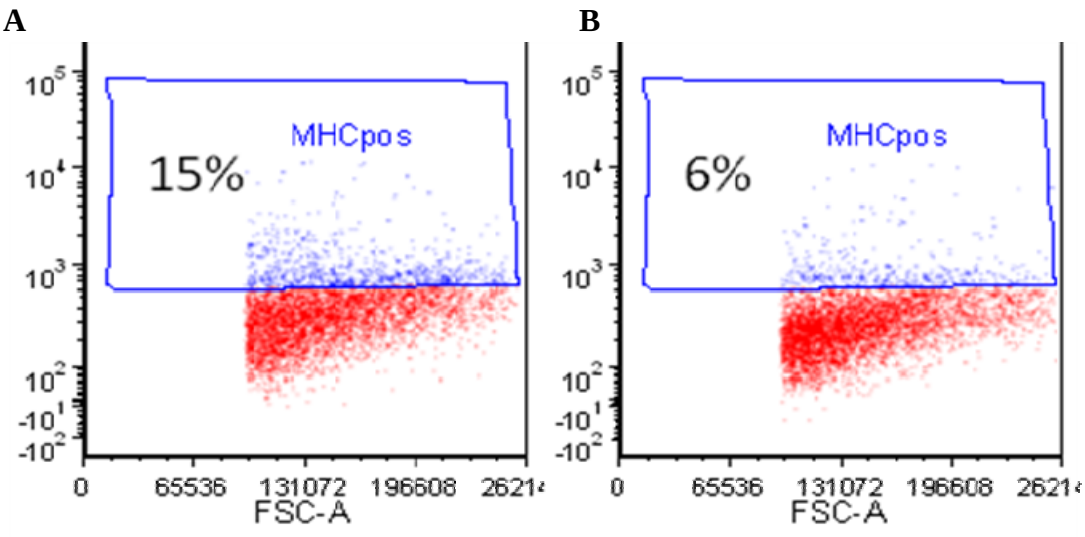

Figure 23, Comparison of cells transfected with either the Amaxa or Neon electroporation systems.
A) Differentiated CMs nucleofected with MHC MB. B) Cells from the same differentiation protocol transfected with MHC MB using the Neon transfection system.

When we transfected the NPPA MBs identified as potential matches by the positive control cell line assay, we were excited to find that the beacons that demonstrated high specificity in the immortalized cell line also appeared to be sensitive to cardiomyocyte differentiated from stem cells. Upon further testing, however, we were not able to reliably reproduce our results. We optimized our procedure for different days, finding that old CMs were easier to transfect with the more stringent 1500/30/2 protocol, while day 14 CMs showed higher efficiency with the 1400/20/2 protocol. We also varied the time that beacons were incubated in cells before analysis, the temperature that they were incubated in, the centrifugation speed that they were rinsed at, and the media that

they were incubated in both after transfection and during flow cytometry. We were not able to see a significant effect due to any of these variables.

To further troubleshoot this issue we compared Neon transfection to the gold standard nucleofection approach developed in previous aims. We found that the most important variables in the protocol were the type of enzyme used for dissociation and number of cells. We also found that while Neon transfection did show similar effects, the fluorescent intensity of MHC MBs was significantly reduced in comparison to nucleofection (Figure 23B) and the results were significantly more variable even when experimental conditions were carefully controlled to reduce the effect of human error. Because of this comparison we decided to proceed with nucleofection rather than Neon transfection for further trials.

NPPA MB transfection

We delivered NPPA beacons into differentiated cardiomyocytes using the nucleofection protocol described in previous aims. We saw that there were significantly fewer NPPA positive cells than MHC positive cells. In addition the number of cells detected as NPPA or MHC positive varied directly with the observed number of beating cells in culture. This number was also quantified using α-actinin staining after the third experiment. To quantify the difference between MHC and NPPA MB positive cells among different cohorts of differentiated CMs, we normalized the beacon positive cells to the MHC MB^{+} cells from the same cohort. We found that the number of MHC+ cells was within 5% of the number of α-actinin positive cells in our control experiment. More importantly, the number of NPPA+ cells is consistently 30% of the number of MHC+ cells. This is a promising sign that the MBs are functioning appropriately and that the

NPPA MBs are selecting a specific subset of cells which varies proportionally with the number of differentiated CMs in culture.

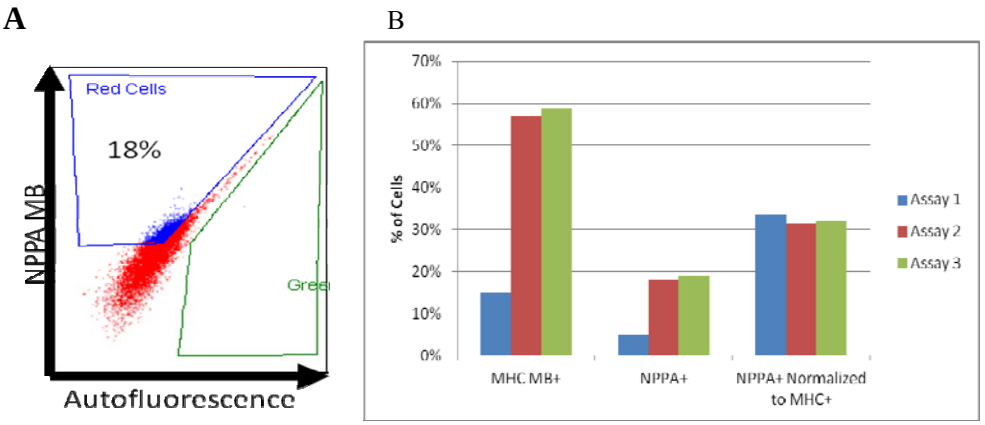

Figure 24, NPPA+ cells in a differentiating culture vary based on the efficiency of CM differentiation A) NPPA positive cells in a differentiing culture. B) NPPA+ cells normalized to the number of CMs in 3 cohorts

In conclusion, we have developed both a more accurate system for quickly screening MBs as well as MBs specific for an important subset of CM cells. Stably transfected cells used as a positive control system introduce an additional level of experimental rigor to molecular beacon specificity testing due to the reduction in dynamic variable in comparison to differentiating or metabolically altered cells. The higher success rates that we saw based on results from the positive control cell line significantly reduced the cost and time associated with beacon specificity testing. IN

addition, this could allow a wider distribution of labor, so that labs without stem cell expertise could provide preliminary data on the usefulness of a specific molecular beacon before providing it to a stem cell lab.

The working CMs that we have isolated using NPPA MBs will be extremely useful to researchers interested in deconvolving the effects of nodal and working cells on either implantation or drug testing. We expect to see a significant reduction in arrythmias and increased conductance when implanting working cells into live hearts. In addition, if both NPPA and MHC beacons are used to isolate CMs we would expect to see a further reduction in the number of undesired cell types present in a cell population. This would reduce the chance of tumorigenesis arising from the implantation of stem cells into tissue (*136*). During drug testing working CMs will display more uniform characteristics of both conduction and contraction, allowing the development of more sensitive assays(*114*). In addition the use of non-genetic isolation methods will enable the development of methods directly applicable to large numbers of iPS cells differentiated in parallel, expanding the information which can be gleaned from preclinical assays (*80*)..

CHAPTER 6

CONCLUSIONS AND FUTURE CHALLENGES

In the three aims of this thesis, the field of cardiac engineering has been advanced by the development of protocols for the isolation of specific subsets of cardiomyocytes. This advance promises to be a significant boon for both cell therapy and drug testing research by increasing control over the inputs of both systems. To publicize these advancements we have published our work in the following research articles:

Wile B, Ban K, Kim S, Park H, Byun J, Cho K, Saafir T, Song M, Yu S, Wagner M, Yoon Y, Bao G. 2013. "Purification of cardiomyocytes from differentiating pluripotent stem cells using molecular beacons targeting cardiomyocyte specific mRNA." *Circulation*. August 2013

Wile B, Ban K, Cho K, Kim S, Song M, Singer J, Syed A, Yu S, Wagner M, Yoon Y, Bao G, "Non-genetic purification of ventricular cardiomyocytes from differentiating embryonic stem cells through molecular beacons targeting a ventricle-specific transcription factor." *In the submission process.*

Wile B, Ban K, Yoon Y, Bao G,"Isolation of high purity cardiomyocytes from differentiating pluripotent stem cells using molecular beacons." Nature Protocols. *Under Review.*

Wile B, Jha R, Xu C, Bao G. "Purification of working cardiomyocytes from differentiation embryonic stem cells based on atrial natriuretic factor messenger RNA." *Manuscript in preparation.*

These articles, in addition to numerous conference posters and presentations, have spawned numerous collaborations and opened avenues to new research questions. We are actively exploring as many of these stories as possible, and we hope that this will grow into an active area of research for years to come.

At the same time, RNA research is becoming more and more widely recognized as an essential field in its own right. The prevalence of transcriptomics as well as the rapid advance of *in situ* RNA probes such as MBs (*127*), mTRIPS (*137*), and nanoflares (*138*) will all provide new and essential tools to advance this area of investigation.

Future opportunities using beacons in developmental biology

We have proven MBs to be a versatile tool which can select various cell types based on abundantly expressed RNA. While this thesis was targeted towards isolating cardiac cells, the opportunity exists to apply the same technique to isolate cell types useful in other areas of developmental biology. Neural, pancreatic, and hepatic developmental biology may benefit from the use of MBs to isolate appropriate cell types.

We have spoken with researchers at Emory and the Joslin Diabetes Center about several of the current roadblocks in neural and pancreatic developmental biology. For neural cell cultures, it is extremely difficult to isolate neurons without contaminating glial cells. Glial cells can significantly affect downstream assays, and they can quickly dominate a neural culture and disrupt measurements (*139*). Likewise, primary pancreatic beta cells can only be purified using an arduous procedure including a sucrose gradient. The resulting mixture of cells can include alpha cells; several research questions are predicated on the interplay between alpha and beta pancreatic cells, so the lack of control in this area hinders sensitive assays (*140*). MBs could be developed to highly expressed

mRNAs in either of these cell types for a more reliable separation method. Currently transfection methods are limiting MBs in this area, but nucleofection has not been tried in either situation, leaving room for improvement.

MBs as a tool to isolate specific cell types

The advent of both iPS cells and gene modification technology point to novel and even more exciting future uses for MBs. MBs are the perfect tool to identify iPS cells or differentiated progeny cells because they do not require genetic modification. Their utility in this regard has already been proven in Chapter 3 of this thesis, however this could be expanded to accelerate the research process with banks of iPS cells used to study specific mutations, such as the LQT1 mutation(*80*). In addition MBs could be used to separate cardiac subtypes to analyze the effect of mutations in different functional capacities.

These mutations could also be induced by gene modification technologies, such as clustered regularly interspaced short palindromic repeats (CRISPR) or Transcription activator-like effector nucleases (TALENs), in standard hESC lines to provide models for research. Currently gene modification technologies suffer from a very low correction rate, making them difficult and expensive to work with because of the sensitivity needed to detect successful modifications(*141*). MBs could be used to select only the modified cell lines, thereby greatly speeding up the process of generating novel cell lines. In addition, MBs could be incorporated into the gene modification process to increase the potency of cell therapies by ensuring that every cell reimplanted into patients has been corrected rather than the current proportions of between 5 and 20%.

APPENDIX A: OLIGONUCLEOTIDE SEQUENCES

Chapter 3:

Beacon Name	Beacon Sequence	Temp (°C)
TNT1	IACCCTCTCTCCATCGGGGATCTTGGGTA	68.1
TNT2	CCCTCTCTTCAGCCAGGCGGTTCTGAGGG	68.4
TNT3	ATCCTCAGCCTTCCTCCTGTTCGAGGAT	64.8
MHC1	CCTCCATCTTCTTCTTCACGGAGG	57.5
MHC2	ITGGCACCAATGTCCCGGCTCTTGCCAA	71.4

Chapter 4:

Name	Sequence	Target	Dye	Quencher
IRX4_1	CACCTAGTTTTGTTATATTAGCCTCCCTAGGTG	AGGGAGGCTAATATAACAAAAC	Cy3	BHQ2
IRX4_2	CCCTGACGTAAACTTTATGCTTCAGGG	CCCTGAAGCATAAAGTTTACGTC	Cy3	BHQ2
IRX4_3	CAGGCAGAGAGTAGAAAGCAGATGCCTG	AGGCATCTGCTTTCTACTCTCTG	Cy3	BHQ2
IRX4_4	CAGGCAGAGAGTAGAAAGCAGATGCCTG	AGGCATCTGCTTTCTACTCTCTG	Alexa488	BHQ1
Control Beacons				
RQ	ACGACGCGACAAGCGCACCGATACGTCGT	GTATCGGTGCGCTTGTCGCG	Cy3	BHQ2
UQ	ACGACGCGACAAGCGCACCGATACGTCGT	GTATCGGTGCGCTTGTCGCG	Cy3	None

Chapter 5:

Probe Name	Probe Sequence	Complex Melting Tm
NPPA_3_UTR_a	CCTCACCCTGCTTGTCCTCCCTGGCTGTTATGTGAGG	71
NPPA_3_UTR_b	ACCTCTTGCAGTCTGTCCCTAGGAGGT	66.3
NPPA_3_UTR_c	CATCACCATGGCAACAAGATGACACAAATGCGTGATG	65.1
NPPA_5_UTR_a	ACCTCTCTTGGCCTACGTCTGTCCCTAGAGGT	71
NPPA_antis_a	ACTGGATCTCTCTGGGCTGGGCTGACTTCCAGT	71.7
NPPALoop150	*CTAGCC*GGGCACGACCTCATCTTCTAAAGGC*TAG*	61.3
NPPALoop190	*CATA*GCTTCTTCATTCGGCTCACTGAGC*TATG*	60.2
NPPA_L680	CACAGTGTTGACAGGAAGCTGCAGCTGTG	66.7
NPPA_L750	CTGCCAATGCATGGGGTGGGAGAGGCAG	69.9
NPPA_antis_a_str	CCTCGATCTCTCTGGGCTGGGCTGACTTCGAGG	70.1
NPPA_antis_b	CATCGAATCCATCAGGTCTGCGTTGGACGATG	67.3
RBMBs		
NPPA_RBMB	[Cy5]CATCGAATCCATCAGGTCTGCGTTGGA CGATGGACGGCAGCGTGCAGCTCTT	67.3
NPPA_RBMB_Stem	[Cy3]GAGCTGCACGCTGCCGTC[BHQ-3]	69.3

REFERENCES

1. V. M. Christoffels, G. J. Smits, A. Kispert, A. F. M. Moorman, *Circulation Research* **106**, 240 (February 5, 2010, 2010).

2. C. Wahlestedt *et al.*, *Proceedings of the National Academy of Sciences* **97**, 5633 (May 9, 2000, 2000).

3. A. K. Chen, O. Davydenko, M. A. Behlke, A. Tsourkas, *Nucleic Acids Research* **38**, (Aug, 2010).

4. R. K. Saiki *et al.*, *Science* **230**, 1350 (Dec 20, 1985).

5. J. C. Alwine *et al.*, *Methods Enzymol* **68**, 220 (1979).

6. M. D. Adams *et al.*, (19920327 DCOM- 19920327).

7. V. E. Velculescu, L. Zhang, B. Vogelstein, K. W. Kinzler, *Science* **270**, 484 (Oct 20, 1995).

8. P. Liang, A. B. Pardee, *Science* **257**, 967 (Aug 14, 1992).

9. M. Schena, D. Shalon, R. W. Davis, P. O. Brown, *Science* **270**, 467 (Oct 20, 1995).

10. G. J. Bassell, C. M. Powers, K. L. Taneja, R. H. Singer, *J Cell Biol* **126**, 863 (Aug, 1994).

11. M. Buongiorno-Nardelli, F. Amaldi, *Nature* **225**, 946 (Mar 7, 1970).

12. S. Behrens, B. M. Fuchs, F. Mueller, R. Amann, *Appl Environ Microbiol* **69**, 4935 (Aug, 2003).

13. A. Tsourkas, M. A. Behlke, S. D. Rose, G. Bao, *Nucleic Acids Res* **31**, 1319 (Feb 15, 2003).

14. G. Bonnet, S. Tyagi, A. Libchaber, F. R. Kramer, *Proc Natl Acad Sci U S A* **96**, 6171 (May 25, 1999).

15. A. Tsourkas, M. A. Behlke, G. Bao, *Nucleic Acids Res* **30**, 4208 (Oct 1, 2002).

16. D. J. States, W. Gish, S. F. Altschul, *Methods* **3**, 66 (1991).

17. S. A. Marras, F. R. Kramer, S. Tyagi, *Nucleic Acids Res* **30**, e122 (Nov 1, 2002).

18. C. J. Yang, W. Lin H Fau - Tan, W. Tan, (20050914 DCOM- 20051128).

19. B. Dubertret, A. J. Calame M Fau - Libchaber, A. J. Libchaber, (20010403 DCOM- 20010816).

20. J. H. Kim, D. Morikis, M. Ozkan, *Sensors and Actuators B-Chemical* **102**, 315 (SEP 13, 2004).

21. Y. Kim, D. Sohn, W. Tan, *Int J Clin Exp Pathol* **1**, 105 (2008).

22. S. A. Kushon *et al.*, *Langmuir* **18**, 7245 (OCT 1, 2002).

23. P. Santangelo, Nitin, N., LaConte, L., Woolums, A., Bao, G., *J Virol* **80**, 682 (2006).

24. P. J. Santangelo, Bao, G., *Nucleic Acids Res* **35**, 3602 (2007).

25. A. Tsourkas, M. A. Behlke, Y. Xu, G. Bao, *Anal Chem* **75**, 3697 (Aug 1, 2003).

26. A. K. Chen, Z. Cheng, M. A. Behlke, A. Tsourkas, *Anal Chem*, (Aug 14, 2008).

27. W. J. Rhee, Santangelo, P.J., Jo, H., Bao, G., *Nucleic Acids Res*, submitted (2007).

28. A. K. Chen, G. Rhee Wj Fau - Bao, A. Bao G Fau - Tsourkas, A. Tsourkas, (20110324 DCOM- 20110705).

29. S. Tyagi, Alsmadi, O., *Biophys J.* **87**, 4153 (2004).

30. A. K. Chen, M. A. Behlke, A. Tsourkas, *Nucleic Acids Research* **35**, (Aug, 2007).

31. A. K. Chen, M. A. Behlke, A. Tsourkas, *Nucleic Acids Research* **36**, (Jul, 2008).

32. M. M. Mhlanga, Vargas, D.Y., Fung, C.W., Kramer, F.R., Tyagi, S., *Nucleic Acids Res* **33**, 1902 (2005).

33. A. K. Chen, M. A. Behlke, A. Tsourkas, *Nucleic Acids Research* **37**, (Dec, 2009).

34. X. H. Peng, Cao, Z. H., Xia, J. T., Carlson, G. W., Lewis, M. M., Wood, W. C., Yang, L., *Cancer Res* **65**, 1909 (2005).

35. N. Nitin, P. J. Santangelo, G. Kim, S. Nie, G. Bao, *Nucleic Acids Res* **32**, e58 (2004).

36. N. C. Price, L. Stevens, *Fundamentals of Enzymology: The Cell and Molecular Biology of Catalytic Proteins*. (Oxford University Press, New York, ed. 3rd, 1999), pp. 478.

37. S. Dokka, Y. Rojanasakul, *Adv Drug Deliv Rev* **44**, 35 (Oct 31, 2000).

38. I. Walev *et al.*, *Proc Natl Acad Sci U S A* **98**, 3185 (Mar 13, 2001).

39. R. V. Giles, C. J. Ruddell, D. G. Spiller, J. A. Green, D. M. Tidd, *Nucleic Acids Res* **23**, 954 (Mar 25, 1995).

40. R. V. Giles *et al.*, *Nucleic Acids Res* **26**, 1567 (Apr 1, 1998).

41. M. A. Barry, A. Eastman, *Arch Biochem Biophys* **300**, 440 (Jan, 1993).

42. E. L. Snyder, Dowdy, S.F., *Curr. Opin. Mol. Ther.* **3**, 147 (2001).

43. J. S. Wadia, Dowdy, S.F., *Curr. Opin. Biotechnol.* **13**, 52 (2002).

44. H. Brooks, Lebleu, B., Vives, E., *Adv Drug Deliv Rev* **57**, 559 (2005).

45. J. S. Wadia, Dowdy, S.F., *Adv Drug Deliv Rev.* **57**, 579 (2005).

46. B. Allinquant *et al.*, *J Cell Biol* **128**, 919 (Mar, 1995).

47. C. M. Troy, D. Derossi, A. Prochiantz, L. A. Greene, M. L. Shelanski, *J Neurosci* **16**, 253 (Jan, 1996).

48. N. Nitin, G. Bao, *Bioconjug Chem* **19**, 2205 (Nov 19, 2008).

49. X. Liu, W. Tan, *Anal Chem* **71**, 5054 (Nov 15, 1999).

50. D. Kambhampati, P. E. Nielsen, W. Knoll, *Biosens Bioelectron* **16**, 1109 (Dec, 2001).

51. F. J. Steemers, J. A. Ferguson, D. R. Walt, *Nat Biotechnol* **18**, 91 (Jan, 2000).

52. H. Kuhn *et al.*, *J Am Chem Soc* **124**, 1097 (Feb 13, 2002).

53. N. Hamaguchi, A. Ellington, M. Stanton, *Anal Biochem* **294**, 126 (Jul 15, 2001).

54. R. Yamamoto, T. Baba, P. K. Kumar, *Genes Cells* **5**, 389 (May, 2000).

55. L. W. van Laake, R. Passier, P. A. Doevendans, C. L. Mummery, *Circ Res* **102**, 1008 (May 9, 2008, 2008).

56. J.-O. Jeong *et al.*, *Circulation Research* **108**, 1340 (May 27, 2011, 2011).

57. R. J. Henning, (20131104).

58. I. Kehat *et al.*, *The Journal of Clinical Investigation* **108**, 407 (2001).

59. M. D. Ungrin, C. Joshi, A. Nica, C. Bauwens, P. W. Zandstra, *PLoS ONE* **3**, e1565 (2008).

60. J. Zhang *et al.*, *Circulation Research* **104**, e30 (2009).

61. M. A. Laflamme *et al.*, *Nat Biotech* **25**, 1015 (2007).

62. S. J. Kattman *et al.*, *Cell Stem Cell* **8**, 228 (2011).

63. T. X. O'Brien, K. J. Lee, K. R. Chien, *Proceedings of the National Academy of Sciences* **90**, 5157 (June 1, 1993, 1993).

64. I. Kehat *et al.*, *The Journal of clinical investigation* **108**, 407 (Aug, 2001).

65. C. Mummery *et al.*, *Circulation* **107**, 2733 (Jun 3, 2003).

66. J. Zhang *et al.*, *Circ Res* **104**, e30 (Feb 27, 2009).

67. T. J. Bartosh, A. A. Wang Z Fau - Rosales, S. D. Rosales Aa Fau - Dimitrijevich, R. S. Dimitrijevich Sd Fau - Roque, R. S. Roque, (20080924 DCOM- 20090105).

68. R. Passier *et al.*, *Stem cells (Dayton, Ohio)* **23**, 772 (Jun-Jul, 2005).

69. Y. Tanabe *et al.*, (20060814 DCOM- 20060913).

70. K. Harris *et al.*, *Toxicological Sciences*, (2013).

71. C. Kane *et al.*, (20140121).

72. A. Martens *et al.*, *Thorac cardiovasc Surg* **60**, PP26 (18.01.2012, 2012).

73. J. Nussbaum *et al.*, *The FASEB Journal* **21**, 1345 (May 1, 2007, 2007).

74. J. C. Moore *et al.*, *Int J Dev Biol* **48**, 47 (Feb, 2004).

75. C. Xu, S. Police, N. Rao, M. K. Carpenter, *Circ Res* **91**, 501 (Sep 20, 2002).

76. L. Yang *et al.*, *Nature* **453**, 524 (May 22, 2008).

77. N. C. Dubois *et al.*, *Nat Biotech* **29**, 1011 (2011).

78. D. A. Elliott *et al.*, *Nat Meth* **8**, 1037 (2011).

79. T. J. Kamp, *Nat Meth* **8**, 1013 (2011).

80. D. Sinnecker *et al.*, *Journal of cardiovascular translational research* **6**, 31 (2013).

81. Y. Yoshida, S. Yamanaka, *Journal of Molecular and Cellular Cardiology* **50**, 327 (Feb, 2011).

82. L. Bu *et al.*, *Nature* **460**, 113 (2009).

83. I. Huber *et al.*, *Faseb J* **21**, 2551 (Aug, 2007).

84. F. Hattori *et al.*, *Nat Meth* **7**, 61 (2010).

85. D. Van Hoof *et al.*, *Journal of Proteome Research* **9**, 1610 (2010/03/05, 2010).

86. H. Uosaki *et al.*, *PLoS ONE* **6**, e23657 (2011).

87. T. Thum *et al.*, *Circulation* **116**, 258 (July 17, 2007, 2007).

88. S. Djurovic, N. Iversen, S. Jeansson, F. Hoover, G. Christensen, *Molecular Biotechnology* **28**, 21 (2004).

89. N. Xu, T. Papagiannakopoulos, G. Pan, J. A. Thomson, K. S. Kosik, *Cell* **137**, 647 (May 15, 2009).

90. G. A. Garinis *et al.*, *Nat Cell Biol* **11**, 604 (2009).

91. K. Terai *et al.*, *Molecular and Cellular Biology* **25**, 9554 (November 1, 2005, 2005).

92. M. A. Frias, M. C. Rebsamen, C. Gerber-Wicht, U. Lang, *Cardiovascular Research* **73**, 57 (January 1, 2007, 2007).

93. O. Caspi *et al.*, *J Am Coll Cardiol* **50**, 1884 (November 6, 2007, 2007).

94. N. Klauke, G. Smith, J. M. Cooper, *Analytical Chemistry* **82**, 585 (2010/01/15, 2009).

95. I.-H. Park *et al.*, *Cell* **134**, 877 (2008).

96. I.-H. Park *et al.*, *Nature* **451**, 141 (2008).

97. A. S. Go *et al.*, *Circulation* **127**, e6 (January 1, 2013, 2013).

98. M. A. Laflamme, C. E. Murry, *Nature* **473**, 326 (2011).

99. K. R. Chien, I. J. Domian, K. K. Parker, *Science* **322**, 1494 (December 5, 2008, 2008).

100. I. J. Domian *et al.*, *Science* **326**, 426 (October 16, 2009, 2009).

101. J.-D. Fu *et al.*, *PLoS ONE* **6**, e27417 (2011).

102. L. Yang *et al.*, *Nature* **453**, 524 (2008).

103. S. Y. Ng, C. K. Wong, S. Y. Tsang, *American Journal of Physiology - Cell Physiology* **299**, C1234 (December 1, 2010, 2010).

104. S.-Y. Liao *et al.*, *Heart Rhythm* **7**, 1852 (2010).

105. M. Muller *et al.*, *Faseb J* **14**, 2540 (Dec, 2000).

106. L. A. Lowe, S. Yamada, M. R. Kuehn, *Development* **128**, 1831 (May, 2001).

107. M. Y. Lee *et al.*, *Stem Cell Research* **8**, 49 (2012).

108. Z.-Z. Bao, B. G. Bruneau, J. G. Seidman, C. E. Seidman, C. L. Cepko, *Science* **283**, 1161 (February 19, 1999, 1999).

109. B. G. Bruneau *et al.*, *Developmental Biology* **217**, 266 (2000).

110. G. F. Wang, W. Nikovits, Z.-Z. Bao, F. E. Stockdale, *Journal of Biological Chemistry* **276**, 28835 (August 3, 2001, 2001).

111. M. Zuker, *Nucleic Acids Res* **31**, 3406 (July 1, 2003, 2003).

112. M. Popenda *et al.*, *Nucleic Acids Res* **40**, e112 (August 1, 2012, 2012).

113. J. R. Hume, A. Uehara, *The Journal of Physiology* **368**, 525 (November 1, 1985, 1985).

114. A. Kuzmenkin *et al.*, *The FASEB Journal* **23**, 4168 (December 1, 2009, 2009).

115. S. V. Raman, G. E. Cooke, P. F. Binkley, *Circulation* **107**, e195 (May 27, 2003, 2003).

116. K. Takahashi *et al.*, *Cell* **131**, 861 (Nov 30, 2007).

117. K. Takahashi, S. Yamanaka, *Cell* **126**, 663 (Aug 25, 2006).

118. F. Hattori *et al.*, *Nat Methods* **7**, 61 (Jan, 2010).

119. A. Bizy *et al.*, *Stem Cell Research* **11**, 1335 (11//, 2013).

120. Q. Zhang *et al.*, *Cell Res* **21**, 579 (2011).

121. W. Rottbauer *et al.*, *Circ Res* **99**, 323 (August 4, 2006, 2006).

122. F. Sheikh *et al.*, *The Journal of clinical investigation* **122**, 1209 (Apr 2, 2012).

123. D. O. Nelson, D. X. Jin, K. M. Downs, T. J. Kamp, G. E. Lyons, *Developmental Dynamics*, n/a (2013).

124. W. J. Rhee, G. Bao, *Bmc Biotechnology* **9**, (Apr, 2009).

125. W. J. Rhee, P. J. Santangelo, H. Jo, G. Bao, *Nucleic Acids Research* **36**, e30 (March 1, 2008, 2008).

126. P. Santangelo, N. Nitin, G. Bao, *Ann Biomed Eng* **34**, 39 (Jan, 2006).

127. K. Ban *et al.*, *Circulation* **128**, 1897 (2013).

128. A. Moretti *et al.*, *New England Journal of Medicine* **363**, 1397 (2010).

129. I. Itzhaki *et al.*, *Nature* **471**, 225 (Mar 10, 2011).

130. P. Liang *et al.*, *Circulation* **127**, 1677 (April 23, 2013, 2013).

131. S. Tohyama *et al.*, *Cell Stem Cell*, (2013).

132. C. Xu *et al.*, *Regenerative medicine* **6**, 53 (Jan, 2011).

133. A. C. Houweling, M. M. van Borren, A. F. Moorman, V. M. Christoffels, *Cardiovascular Research* **67**, 583 (Sep 1, 2005).

134. T. Horsthuis *et al.*, *Circ Res* **105**, 61 (Jul 2, 2009).

135. A. C. Houweling, S. Somi, M. J. Van Den Hoff, A. F. Moorman, V. M. Christoffels, *Anat Rec* **266**, 93 (Feb 1, 2002).

136. A. Behfar *et al.*, *J Exp Med* **204**, 405 (Feb 19, 2007).

137. E. Alonas *et al.*, (20140128).

138. A. E. Prigodich *et al.*, (20120313 DCOM- 20120716).

139. D. E. Willis *et al.*, (20111013 DCOM- 20111206).

140. G. Kilic *et al.*, (20140205).

141. Y. Lin *et al.*, (20140120).

VITA

Brian M. Wile

Brian was born in Snellvile, GA. He attended public schools in Gwinnett County, Georgia, and received a B.E. in Biomedical Engineering from Vanderbilt University, Nashville, Tennessee in 2009 before coming to Georgia Tech to pursue a doctorate in Biomedical Engineering. When he is not working on his research, Brian enjoys attending hiking with his dog, Woodford, and watching The Walking Dead with his beloved girlfriend Jen.