Using Nucleic Acids to Repair β-Globin Gene Mutations

by

Monique Natasha Kierlin-Duncan

University Program in Genetics and Genomics
Duke University

Date: 4/23/07

Approved:

Bruce A. Sullenger Ph.D., Supervisor

Mariano Garcia-Blanco M.D./Ph.D.

Matthias Gromeier M.D.

Marcy Speer Ph.D.

Marilyn Telen M.D.

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University Program in Genetics and Genomics in the Graduate School of Duke University

2007
ABSTRACT

Using Nucleic Acids to Repair β-Globin Gene Mutations

by

Monique Natasha Kierlin-Duncan

University Program in Genetics and Genomics
Duke University

Date: 4/23/07

Approved:

Bruce A. Sullenger Ph.D., Supervisor

Mariano Garcia-Blanco M.D./Ph.D.

Matthias Gromeier M.D.

Marcy Speer Ph.D.

Marilyn Telen M.D.

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University Program in Genetics and Genomics in the Graduate School of Duke University

2007
Abstract

Nucleic acids are an emerging class of therapeutics with the capacity to repair both DNA and RNA mutations in clinically relevant targets. We have used two approaches, mobile group II introns and Spliceosome Mediated RNA Trans-splicing (SMaRT), to correct β-globin mutations at the DNA and RNA levels respectively. We show that the group II intron inserts site-specifically into its DNA target, even when similar targets are available. Experiments transitioning this therapeutic into mammalian cell systems are then described. We also illustrate how SMaRT RNA repair can be used to correct β-globin mutations involved in sickle cell disease and some forms of β-thalassemia. We uncovered diverse repair efficiencies when targeting sickle cell versus β-thalassemia transcripts in mammalian cells. Possible reasons for this and how it might direct target choice for the SMaRT therapeutic approach are both discussed. The therapeutic molecule in SMaRT, a Pre-Trans-splicing Molecule or PTM, is also delivered via lentivirus to erythrocyte precursors cultured from the peripheral blood of sickle cell patients. Preliminary results from these experiments are discussed.
Dedication

I’d like to dedicate my doctoral work and this dissertation to my mom, Stefannie Valencia-Kierlin, who passed away without knowing I went to graduate school. She would have been a wonderful source of love and encouragement.
Table of Contents

Abstract ........................................................................................................ iv
Dedication ....................................................................................................... v
Table of Contents ........................................................................................ vii
List of Tables ................................................................................................ ix
List of Figures ................................................................................................ x
Acknowledgements ...................................................................................... xiii
Chapter 1: Introduction ............................................................................. 1
  1.1 The β-globin Gene ................................................................................. 1
  1.2 Sickle Cell Disease and β-Thalassemia ................................................... 4
  1.3 Therapeutic Approaches to Repair DNA and RNA ................................. 8
Chapter 2: DNA Repair Using the Lactococcus lactis Group II Intron .......... 14
  2.1 Introduction to the L. lactis Group II Intron ............................................ 14
  2.2 Testing the Specificity of Retrohoming in a Bacterial System ................. 20
    2.2.1 Results .......................................................................................... 20
  2.3 Moving the L. lactis Group II Intron into a Mammalian System ............... 30
    2.3.1 Results .......................................................................................... 30
  2.4 Conclusions from Sections 2.2 and 2.3 .................................................. 42
  2.5 Materials and Methods for Sections 2.2 and 2.3 .................................... 44
Chapter 3: RNA Repair Using Spliceosome-Mediated RNA Trans-Splicing (SMaRT) Technology ......................................................... 50
3.1 Introduction to SMaRT Technology.................................................................50

3.2 Using 5’PTMs to repair the sickle cell mutation in mammalian cell culture
   experiments ........................................................................................................58
   3.2.1 Summary .................................................................................................58
   3.2.2 Introduction ............................................................................................58
   3.2.3 Results ....................................................................................................59
   3.2.4 Discussion ...............................................................................................67
   3.2.5 Materials and Methods ..........................................................................69

3.3 Using 5’PTMs to repair the sickle cell mutation in primary proerythroblasts from sickle cell patients.................................................................73
   3.3.1 Summary .................................................................................................73
   3.3.2 Introduction ............................................................................................73
   3.3.3 Results ....................................................................................................75
   3.3.4 Discussion ...............................................................................................104
   3.3.5 Materials and Methods ..........................................................................107

3.4 Using 5’PTMs to repair β-thalassemia mutations ........................................118
   3.4.1 Summary .................................................................................................118
   3.4.2 Introduction ............................................................................................118
   3.4.3 Results ....................................................................................................122
   3.4.4 Discussion ...............................................................................................136
   3.4.5 Materials and Methods ..........................................................................140
Chapter 4: Conclusions and Future Work ................................................................. 142

4.1 DNA Repair Using the L. lactis Group II Intron.................................................. 142

4.1.1 The L. lactis Group II Intron’s Target Specificity Results in a
Potentially Safer DNA Repair Therapeutic............................................................. 142

4.1.2 Heterochromatin and the Mammalian Cell Environment May Pose
Challenges to This Therapeutic Approach............................................................. 143

4.2 RNA Repair Using 5’PTMs and SMaRT Technology......................................... 144

4.2.1 The Potential Impact of Nuclear Double-Stranded RNA Processes on
SMaRT ..................................................................................................................... 144

4.2.2 Target Choice and SMaRT .............................................................................. 150

References.................................................................................................................. 152

Biography ................................................................................................................. 168
List of Tables

Table 2.1 Repair efficiency of the L1.LtrB group II intron using an *E. coli* assay ..........25

Table 2.2 Target mutants constructed from various genomic sequences then used with pIntronLac in a *lacZ E. coli* mobility assay. .................................................................27

Table 2.3 The wildtype and mutant β-globin DNA targets. ........................................29

Table 3.1 The four β-thalassemia mutations used in this study.................................121
List of Figures

Figure 1.1 The chromosome 11 β-globin gene cluster and two forms of hemoglobin........2

Figure 2.1 The mobile L. lactis LtrB group II intron.....................................................16

Figure 2.2 Excision and retrohoming mechanisms of the mobile L. lactis LtrB group II intron.................................................................17

Figure 2.3 Sequence requirements for retargeting the Ll.LtrB to novel sites..............19

Figure 2.4 Ll.LtrB retargeted to bind a therapeutically relevant DNA target. ............20

Figure 2.5 The lacZ mobility assay in E. coli.................................................................22

Figure 2.6 Assay for detecting Ll.LtrB intron forward splicing in mammalian cells.....31

Figure 2.7 Resulting protected RNA fragments (probe binding sites and gel schematics) after using either the “react” or “unreact” RPA probes in the presence/absence of Ll.LtrB forward splicing in mammalian cells.............................................32

Figure 2.8 Testing the react and unreact controls for the RPA experiment....................33

Figure 2.9 Using different amounts of react probe to detect Ll.LtrB forward splicing in mammalian cells..........................................................................................................................35

Figure 2.10 The group II Ll.LtrB transcript produced in mammalian cells will forward splice in vitro under high salt conditions.................................................................38

Figure 2.11 Providing the hLtrA protein in cis..............................................................40

Figure 2.12 Western blot illustrating low levels of hLtrA protein production in mammalian cells..........................................................................................................................41

Figure 3.1 Exon replacement via trans-splicing using Spliceosome-Mediated RNA Trans-Splicing Technology (SMaRT).................................................................52

Figure 3.2. 5’PTM constructs. .....................................................................................61

Figure 3.3. Sickle β-globin transcripts are corrected via 5’PTM trans-splicing. .........62

Figure 3.4 Quantifying the levels of sickle β-globin repair........................................65
Figure 3.5 Possible RNA secondary structure of the 5’PTM’s tag.................................68
Figure 3.6 Erythropoiesis and the proerythroblast lineage........................................74
Figure 3.7 CD34+ cells in sickle cell patient peripheral blood can be cultured to proerythroblasts in vitro.................................................................76
Figure 3.8 The day 10 cultured cells are primarily erythrocyte precursors......................77
Figure 3.9 The 5’PTM1 and β-globin transcripts are both expressed in proerythroblasts cultured from sickle cell patients after transfection or electroporation of 5’PTM1EGFPN1.................................................................79
Figure 3.10 5’PTM1 repairs the sickle β-globin transcript in patient proerythroblasts after transfection or electroporation of 5’PTM1EGFPN1..................................................80
Figure 3.11 Lentiviral constructs.....................................................................................84
Figure 3.12 Lentiviral transduction scheme I.................................................................85
Figure 3.13 The vsv-g GFP lentivirus will infect primary proerythroblasts from sickle cell patients........................................................................................................87
Figure 3.14 The 5’PTM1 transcript is expressed from vsv-g or amphotropic enveloped lentiviruses within sickle cell patient proerythroblasts........................................87
Figure 3.15 Lentiviral delivery of 5’PTM1 into sickle cell patient proerythroblasts results in repair of the sickle cell transcript.................................................................90
Figure 3.16 RNase Protection Assay..............................................................................93
Figure 3.17 Possible false positive result in the RPA experiment when using high levels of input RNA.................................................................94
Figure 3.18 Lentiviral transduction scheme II.................................................................96
Figure 3.19. Viral transduction scheme II leads to high lentiviral GFP expression in proerythroblasts and the viral envelope may have an impact on transduction efficiency.................................................................98
Figure 3.20 The 5’PTM1 transcript is expressed from vsv-g or amphotropic enveloped lentiviruses within sickle cell patient proerythroblasts using viral transduction scheme II.................................................................102
Figure 3.21 5’PTM1 lentiviral delivery into sickle cell patient proerythroblasts using transduction scheme II results in repair of the sickle cell transcript. .......................... 103

Figure 3.22 The four β-thalassemia mutants are corrected via 5’PTM-driven trans-splicing. ........................................................................................................ 123

Figure 3.23 5’PTMs can restore the normal splice product in each β-thalassemia mutant. ....................................................................................................................... 126

Figure 3.24 Testing a HaeIII restriction enzyme assay for potential use in quantifying β-thalassemia repair levels. .................................................................................. 130

Figure 3.25 Quantification of repair levels engendered by each 5’PTM targeting the four β-thalassemia mutants ................................................................................................. 133

Figure 3.26 Percent repair of β-thalassemia mutants using various 5’PTM constructs. 135

Figure 4.1 Hypothetical PTM/target pre-mRNA dsRNA for use in Drosha activity testing. ................................................................................................................................... 148
Acknowledgements

This is going to take two pages and if this were the Academy Awards, I would definitely hear the background music pushing me off stage. It’s worth it though.

I would first like to thank my thesis committee members, Matthias Gromeier, Mariano Garcia-Blanco, Marilyn Telen, Marcy Speer, and past member Andy Amalfitano for their guidance during my doctoral work. Matthias, your practical advice during the committee meetings was always helpful and thanks also for bestowing upon me the title of “Honorary Gromeier Lab Member for Social Purposes”. Mariano and Marcy, I remember interviewing with both of you 6 years ago. Mariano, our lengthy discussion regarding the selective advantage of alternative splicing was an unexpected and wonderful interview topic. Marcy, your enthusiasm for the UPGG program was so genuine and your ebullient personality put me at ease during that first meeting. Dr. Telen, thank you for ushering me through numerous clinical realms during my primary cell work. Thank you also for introducing me to Dr. Murat Arcasoy who proved to be an extremely valuable resource to me. Andy, thank you for offering insight into the therapeutic aspects of my project.

Bruce, thank you for allowing me to join your lab and for all your support. Thanks especially for giving me the freedom to direct the course of my graduate work and for trusting me to try things on my own. Your collaborative spirit was also infectious and I appreciate what I’ve learned from you in that respect. Thank you also for introducing me to Dr. John Gray at St. Jude. Your introduction resulted not only in a
fruitful scientific collaboration, but also in later opportunities to meet with Drs. Derek Persons and Art Nienhuis while in Memphis.

Dad and Lara, thanks for asking how my research was going and encouraging me when things weren’t working. Dad, I’ll always remember your phone calls beginning with, “how are the amoebas growing?” which had absolutely nothing to do with my thesis topic, but you would ask almost every time. I appreciate the effort.

Brent, thanks for being such a supportive husband, especially while I was writing this thesis. I felt absolutely no pressure to shower, clean anything, or avoid overspending at Starbucks from you and I totally appreciate that. Thanks for taking Dante and Alighieri on their walks when I was busy and for just understanding how hard graduate school can be. I love being a Kierlin-Duncan.

That’s it, I guess. I should probably thank my green sweatpants that I wore for one month straight while I wrote this, but that would just be strange. (thanks green sweatpants…)
Chapter 1: Introduction

1.1 The β-globin Gene

The β-globin gene product, the β-chain, is a vital member of the oxygen transport pathway in the bloodstream. Two β and two α chains comprise the core structure of hemoglobin, which is the tetrameric molecule responsible for delivering oxygen to tissues throughout the body (Perutz, 1960). β-chains are found in the adult form of hemoglobin, but two other hemoglobin types are made during development: embryonic and fetal. Embryonic and fetal hemoglobin carry unique β-like chains called ε and γ-chains respectively. The genes encoding for these β-like chains are located near the β-globin gene in the chromosome 11 β-globin gene cluster. Here, 5 genes, ε, γG, γA, δ and β are transcriptionally regulated by both proximal elements and by an upstream regulatory region called the locus control region to produce the different β-like chains throughout development (figure 1.1).
Figure 1.1 The chromosome 11 β-globin gene cluster and two forms of hemoglobin

A. The five primary β-globin-like genes are shown above (the gene in the white box is a pseudo-gene) and their developmental expression is both proximally and distally controlled (by the locus control region, not shown here). The β-globin gene is described in detail below and contains three exons. Its gene product, the β-chain, is a vital component of adult hemoglobin. B. The fetal and adult forms of hemoglobin. Each chain in the tetrameric molecule associates with an iron-containing heme component (in yellow). The α, β, and γ-chains are in gray, red, and blue respectively.
Together with α-type chains generated from chromosome 16, various β-type chains are temporally expressed to produce the different forms of hemoglobin in a process known as hemoglobin switching (Nienhuis & Stamatoyannopoulos, 1978). The molecular basis for hemoglobin switching involves both competition for interaction with the locus control region (LCR) and independent regulation of each of the genes on the β-globin gene cluster (“competitive” and “autonomous”, reviewed in (Stamatoyannopoulos, 2005)). The precise nature of the LCR interaction is under debate, but one model, the facilitated chromatin-looping model posed by Li and colleagues, allows for a combination of various hypotheses under a generalized looping model (Li et al., 2006). The original looping hypothesis proposed by Wang and colleagues suggested that the LCR and local promoter interact directly thus causing the intervening sequence to form a loop-like structure (Wang & Giaever, 1988). Li and colleagues postulated that perhaps histone modifications stimulate changes in the flexibility of chromatin and this then influences how the “loops” form. The histone modifications themselves might be regulated by elements described within other theories of LCR interaction, namely the “linking” and “tracking” hypotheses that suggest that certain proteins either provide the link between the LCR and promoter or that LCR-related proteins track along the DNA strand until they encounter a promoter (Tuan et al., 1992; Bulger & Groudine, 1999).

Control over the chromosome 11 β-globin gene cluster is of interest not only because of its complex regulatory mechanisms, but for potential therapeutic motivations
as well. Fetal hemoglobin exerts a protective effect against polymerization of adult hemoglobin in sickle cell anemia and the γ-chain may be a useful substitute for defective β-chain synthesis in β-thalassemia (Nagel et al., 1979; Perrine et al., 1993). These approaches will be explored further in the next section.

1.2 Sickle Cell Disease and β-Thalassemia

Sickle cell disease (SCD) and β-thalassemia both originate from mutations in the β-globin gene. Although the two hematological conditions share some traits like hemolytic anemia, other facets are quite distinct. SCD is a hemoglobinopathy (a structural aberration in the hemoglobin molecule) whereas β-thalassemia generally refers to a reduction in the level of normal β-chains produced (although some dominant negative mutations have been documented (Fei et al., 1989)). SCD is primarily caused by one of two different mutations, an A to T base substitution in β-globin’s 6th codon (GAG glutamic acid to GTG valine) or a G to A base substitution in the same codon (GAG glutamic acid to AAG lysine). Individuals homozygous for the first mutation have hemoglobin SS disease, commonly known as sickle cell anemia. The second allele is usually found in patients with hemoglobin SC disease who are compound heterozygotes for the two mutations. Conversely, β-thalassemia can be caused by a broad assortment of mutations in the β-globin gene and hence it is argued that this condition also results in a more heterogeneous collection of possible clinical phenotypes (Thein, 2005a). This sentiment is also suggested by the common use of sub-classifications such as β0 or β+.
thalassemia to reflect the spectrum of clinical severity resulting from varying levels of β-chain production.

However, overlap does exist among the phenotypic characteristics of SCD and β-thalassemia. Since both diseases are hemolytic anemias, the patient populations share similar clinical concerns and treatment options for this problem. Transfusion therapy may be indicated in either diagnosis and if used, similar surveillance measures for potential deleterious side effects like increases in blood viscosity or iron levels must be taken (Kwiatkowski & Cohen, 2004; Swerdlow, 2006). Both patient populations may also be given hydroxyurea to boost fetal hemoglobin levels. A subset of patients even present with both conditions simultaneously, which is caused by compound heterozygosity at the β-globin locus (HbSpthal).

SCD was first described in 1910 and since then, we have gained much knowledge regarding its molecular, physiological, clinical, and epidemiological aspects (Herrick, 1910; Scriver, 1930). The clinical symptoms include anemia, acute chest syndrome, painful crises, stroke, and pulmonary hypertension among others (see review (Stuart & Nagel, 2004)). The glutamic acid to valine amino acid substitution at codon 6 introduces a hydrophobic residue that binds within a hydrophobic pocket formed in another hemoglobin molecule during oxidative stress (Pauling et al., 1949; Ingram, 1956; Wishner et al., 1975). This leads to intra-cellular polymerization of hemoglobin tetramers and rigid deformation/destruction of the red blood cell. The resulting pathophysiological sequelae include micro-vascular occlusion (due to a number of factors causing slow
movement of sickle cells through the vasculature), localized adhesion to endothelial cells, inflammation, anemia, hypoxia and related tissue damage (Hoover et al., 1979; Pearson et al., 1985; Kaul et al., 1989; Gee & Platt, 1995; Vichinsky et al., 2000; Bookchin & Lew, 2002; Kaul & Fabry, 2004; Belcher et al., 2006).

The mild clinical phenotype observed in sickle cell patients with hereditary persistence of fetal hemoglobin (HPFH) was an early indicator of fetal hemoglobin’s protective effect in SCD (Perrine et al., 1972; Stamatoyannopoulos et al., 1975; Sunshine et al., 1978; Powars et al., 1984). Studies found that fetal hemoglobin levels greater than 20% correlated with a marked reduction in symptoms in sickle cell patients and that even 10% provided some protection (Powars et al., 1984). Fetal hemoglobin exerts its protective effect by binding to sickle hemoglobin and disrupting the polymerization event (Bookchin et al., 1975; Nagel et al., 1979). A number of approaches have been adopted to increase fetal hemoglobin levels in sickle cell patients with the most common being administration of hydroxyurea (Platt et al., 1984; Charache et al., 1995; Ware et al., 2002; Steinberg et al., 2003; Hankins et al., 2005). Hydroxyurea is a chemotherapeutic agent that inhibits ribonucleotide reductase and eventually leads to cell death. It is thought that its cytotoxicity causes an altered differentiation program (more rapid) within bone marrow that results in increased fetal hemoglobin production (Letvin et al., 1984; Veith et al., 1985). However, this is not the sole mechanism posed by researchers. Hydroxyurea may also stimulate downstream γ-globin gene transcription by its in vivo oxidation to nitric oxide or by inducing expression of a particular small GTP-binding protein. Both are
hypothesized to trigger a signaling cascade leading to increased expression of the γ-globin gene (Cokic et al., 2003; Tang et al., 2005). Other pharmacological and gene therapy approaches have targeted γ-globin gene expression as a method for increasing fetal hemoglobin and these will be discussed in the following review of β-thalassemia.

β-thalassemia is caused by an imbalance in the α to β chain ratio in erythrocyte precursors. The clinical severity often depends on the levels of normal β-chain production and/or the presence of deleterious β-chains. Anemia is caused by either destructive inclusion body formation during early erythropoiesis or via erythrocyte damage in the microvasculature (review in (Thein, 2005a)). The higher relative numbers of α-chains resulting from decreased β-chain production leads to aggregation of the α-chains and formation of inclusion bodies (Schrier, 2002). Erythrocyte precursors are often destroyed as a consequence of intracellular inclusion body accumulation and ineffective erythropoiesis results. Mature thalassemic erythrocytes are also susceptible to damage while traversing the spleen, which causes not only further anemia, but injury to splenic tissues as well due to localized cellular aggregation (Schrier, 2002).

Therapies for β-thalassemia focus primarily on restoring the α to β chain ratio. Pharmacological agents that have been used in sickle cell anemia to increase levels of fetal hemoglobin such as hydroxyurea, 5-azacytidine, and arginine butyrate have also been successfully applied to patients with β-thalassemia (review in (Perrine, 2005)). The γ-globin chains bind the free α-chains and hence prevent formation of inclusion bodies and provide a source of functional hemoglobin (review in (Thein, 2005b)). The potential
mechanisms for hydroxyurea were discussed previously, but its widespread use is likely due to its combined efficacy and ease of administration. 5-azacytidine inhibits DNA methylation and thus prevents normal γ-globin gene silencing in adult erythropoiesis (Ley et al., 1983a; Ley et al., 1983b). However, carcinogenic concerns now prohibit its use in the U.S. (Fathallah & Atweh, 2006). Butyrate protects the upstream region of γ-globin from deacytylation and thus maintains a transcriptionally favorable open chromatin conformation (Perrine et al., 1989; Perrine et al., 1993). Low erythropoietin production is also sometimes found in patients with β-thalassemia and co-administration of this growth factor might be useful in conjunction with any of these approaches (Galanello et al., 1994).

### 1.3 Therapeutic Approaches to Repair DNA and RNA

The hematological diseases described above have been targets for both DNA and RNA-based therapeutics including lentiviral, ribozyme, trans-splicing, and antisense oligonucleotide approaches. Since mobile group II introns (for DNA repair) and SMaRT technology (for RNA repair) will be discussed in detail in the following chapters, they will not be covered in this particular section.

Lentiviral gene therapy for sickle cell anemia and β-thalassemia invokes the lentivirus’ unique ability to integrate into cells that are not actively dividing (Lewis & Emerman, 1994). This non-dividing state is often found in hematopoietic stem cells (HSC) and lentiviruses have been shown to effectively deliver transgenes into this cell
type (Reiser et al., 1996; Chen et al., 2000; May et al., 2000; Pawliuk et al., 2001; Rivella et al., 2003; Vacek et al., 2003).

Work by May and colleagues used a lentiviral vector to deliver the chromosome 11 LCR and complete wild type β-globin gene into normal and β-thalassemia mouse models (May et al., 2000). They showed long-term integration of the transgene by performing secondary bone marrow transplants using marrow from donor mice that were 24 weeks post-transplant and assaying the recipient mice 12 weeks after this. Transgene expression was also detected in the primary transplant mice at 24 weeks post-treatment using one particular viral vector.

Since γ-globin gene expression is desirable in both sickle cell anemia and β-thalassemia, investigators have since incorporated either γ-globin genes or γ-like modified β-globin genes into their lentiviral constructs (McCune et al., 1994; Levasseur et al., 2003; Hanawa et al., 2004; Levasseur et al., 2004). One motivation for maintaining a β-globin “backbone” in the modified β-globin approach is to avoid the need for γ-specific transcription factors that may be absent during adult erythropoiesis while still providing the functionally desirable characteristics of fetal hemoglobin (Levasseur et al., 2004). Introducing a transgene similar to the normally active globin gene in adult bone marrow may not only improve transcription, but it may also result in long-term therapeutic expression. Persons and colleagues have used a γ-globin gene with an in vivo selectable marker as another approach for sustaining transgene expression over time (Persons et al., 2004).
Safety concerns surrounding retroviral gene therapy are currently focused on off-target insertional events into genomic DNA. The retroviral insertion and subsequent activation of the LMO2 proto-oncogene in two clinical trial patients emphasized the need to address this serious potential outcome (Hacein-Bey-Abina et al., 2003). Since then, investigators have examined the integrating preferences of both onco-retroviruses and lentiviruses (Mitchell et al., 2004). They found that lentiviruses tend to integrate into actively expressing regions whereas onco-retroviruses appear to prefer promoter regions (active or not) (Mitchell et al., 2004). A study in 2004 specifically examined retroviral integration tendencies within hematopoietic stem cells and investigators found integration patterns similar to those described for other cell types (Hematti et al., 2004). Reducing the impact of onco-retroviral insertional mutagenesis, where promoter activation is the primary concern, may come from improvements in viral vector insulator elements. Lentiviral insertional mutagenesis is more likely to result in gene disruption events and hence insulators may not ameliorate this issue. However, since ex vivo correction of HSC with subsequent transplant is possible for many hematological diseases, perhaps screening processes prior to transplant could remove those cells containing such gene disruptions. Investigators are also devising ways to keep overall lentiviral particle levels as low as possible while maintaining efficacy and this too may reduce the risk for deleterious off-target incidents (Levasseur et al., 2003).

RNA repair methodologies are desirable in instances when it is important to preserve the target gene’s endogenous regulatory mechanisms since the therapeutic is
directed to the RNA product instead of genomic DNA. A number of different approaches have been developed for RNA repair of mutant β-globin transcripts including group I ribozymes (Sullenger & Cech, 1994; Lan et al., 1998; Byun et al., 2003), antisense oligonucleotides (Dominski & Kole, 1993; Sierakowska et al., 1996; Gorman et al., 2000; Lacerra et al., 2000; Suwanmanee et al., 2002; Vacek et al., 2003), and siRNA-based methodologies (Dykxhoorn et al., 2006; Samakoglu et al., 2006).

The group I ribozyme from *Tetrahymena thermophila* has been studied in our lab as an approach to repair the sickle cell transcript. Lan and colleagues used a group I ribozyme targeted to an upstream uridine in the βs transcript to replace the downstream mutant RNA with a γ-globin substitute via catalytic RNA trans-splicing (Lan et al., 1998). Further work by Byun et al found that extended a targeting sequence (EGS) and adding an additional interaction within the ribozyme (P10) improved the efficiency of trans-splicing to the sickle globin transcript (Byun et al., 2003).

Ryszard Kole’s group and colleagues have completed a number of studies using antisense oligonucleotides (ASO) to repair β-globin splicing mutants that cause forms of β-thalassemia. They have used ASOs to block spliceosomal usage of cryptic splice sites in β-globin’s intervening sequence 2 (IVS-2) region and shown efficacy in mammalian cells lines as well as in primary cells (Dominski & Kole, 1993; Sierakowska et al., 1996; Gorman et al., 2000; Lacerra et al., 2000; Suwanmanee et al., 2002; Vacek et al., 2003). They have also generated a mouse model to one of these β-thalassemia mutants and
demonstrated the ability of an ASO to direct splicing of a pre-mRNA in vivo (Lewis et al., 1998; Roberts et al., 2006).

RNA silencing and specifically that via the small interfering RNA (siRNA) mechanism has been recently applied to β-globin targets. Samakoglu and colleagues published a study using lentiviral gene delivery of a γ-globin gene carrying a short hairpin RNA (shRNA) within either IVS-2 or the 5’ untranslated region (5’UTR) of the transgene (Samakoglu et al., 2006). They demonstrated targeted knockdown of sickle β-globin expression concurrent with increased γ-globin RNA levels after transduction into MEL (murine erythroleukemia) cells endogenously expressing βs and green fluorescent protein (GFP). They then infected CD34+ cells from sickle cell patients with these constructs and found similar results. Dykxhoorn and colleagues used various siRNAs to determine the mechanism’s ability to distinguish between transcript targets differing by a single nucleotide, as is the case in normal versus sickle β-globin RNA (Dykxhoorn et al., 2006). They found that a mismatch at one particular site is critical to avoid cleavage and that the nature of the mismatch need be a bulky purine:purine interaction for the most protection. The βs siRNA/βwt target pairing at that critical position was purine:purine, whereas it was not for βs siRNA/βs. This therefore suggested that the siRNA could be used for discriminate cleavage of the βs transcript.

The β-globin gene, sickle cell disease and β-thalassemia together provide an excellent canvas for testing a number of clinically applicable therapeutic approaches. The relatively small size of the β-globin gene makes technical aspects such as cloning quite
manageable and the gene has been studied so thoroughly that rational design can be used to manipulate its function. Sickle cell disease and β-thalassemia span a wide spectrum of molecular and clinical pathologies and hence provide an opportunity for the different nucleic acid therapies to find their niche.
Chapter 2: DNA Repair Using the Lactococcus lactis Group II Intron

2.1 Introduction to the L. lactis Group II Intron

*Lactococcus lactis*, a gram-positive bacterium, encodes a group II mobile intron within its LtrB gene. Shearman and colleagues first described this intron and Mills further characterized it as a regulatory element for LtrB gene expression (Mills et al., 1996; Shearman et al., 1996). The LtrB gene product is a relaxase enzyme that is used by *L. lactis* for conjugation and plasmid transfer (Mills et al., 1996). The group II intron’s forward splicing reaction releases it from the gene thus allowing for uninterrupted expression of the LtrB relaxase product.

The *L. lactis* group II intron (hereafter referred to as “Ll.LtrB”) is not only able to remove itself from double-stranded (“ds”) DNA via forward splicing, but it can reverse splice back into DNA as well in a process termed “retrohoming” (Meunier B, 1990; Matsuura et al., 1997; Cousineau et al., 1998). The six domain Ll.LtrB intron contains sequence elements called EBS1, EBS2, and δ that are used for forward splicing and subsequent detection of its intended target DNA (figures 2.1 and 2.2). Matsuura and colleagues devised an assay in *E. coli* whereby Ll.LtrB forward splicing and retrohoming processes could be studied (Matsuura et al., 1997). This system was later used to elucidate detailed information on the intron’s forward splicing mechanism (Saldanha et al., 1999) and targeting rules for the intron’s mobility activities (Mohr et al., 2000). Cousineau and colleagues later presented a derivation of Matsuura’s assay as a method
for in-depth study of the retrohoming reaction (Cousineau et al., 1998). The researchers added a group I self-splicing intron into Ll.LtrB’s domain IV in an effort to determine the role of the Ll.LtrB RNA in retrohoming. They found that the group I sequence was absent after mobility and concluded that since self-splicing had occurred, a portion of the Ll.LtrB existed as an RNA intermediate during retrohoming (Cousineau et al., 1998).

The potential for Ll.LtrB retargeting was another finding within Matsuura’s 1997 publication. A short time prior to this publication, Guo and colleagues had published a study describing the retargeting capabilities of the yeast a12 group II intron (Guo et al., 1997). Matsuura’s group made similar nucleotide substitutions in the Ll.LtrB’s version of the EBS1 sequence and tested the intron’s ability to mobilize into complementary and non-complementary target DNA (Matsuura et al., 1997). They found that, like the yeast group II intron, Ll.LtrB mobility could be manipulated (Matsuura et al., 1997).
Figure 2.1 The mobile *L. lactis* LtrB group II intron.
This figure is taken from Guo et al, Science 2000 (complete citation in References section). The Ll.LtrB intron has six domains with its catalytic core primarily within domain V. Three sequence elements in domain I termed EBS1 (Exon Binding Site 1), EBS2 (Exon Binding Site 2), and δ provide the specificity for selective mobilization into DNA target sites. Domain IV encodes the IEP (Intron Encoded Protein) that in Ll.LtrB is called ltrA. This protein provides maturase, endonuclease and reverse transcriptase capabilities to the mobile intron. It can be moved downstream from Ll.LtrB to provide room for therapeutic cargo in domain IV.
Figure 2.2 Excision and retrohoming mechanisms of the mobile *L. lactis* LtrB group II intron.
The intron sequence is first transcribed from its DNA template and the intron-encoded-protein (IEP here is ltrA) is translated. The protein binds to the intron RNA sequence and assists this catalytic RNA in performing a forward splicing reaction that results in excision from (and ligation of) the remaining transcript. The functional unit of LLLtrB is a ribonucleoprotein (RNP) that now mobilizes to its target DNA site. The IEP binds E1 of the target site and this binding allows the RNA portion of the RNP to basepair with the DNA target via its EBS2, EBS1, and δ sequences. The RNA then cleaves the sense strand of the target DNA and reverse splices into this strand. The IEP cleaves the antisense strand and acts as a reverse transcriptase to make the first DNA strand, using the intron RNA as a template. Host mechanisms are thought to then repair the gap, degrade the intron RNA, synthesize the second DNA strand and repair this final gap.
Using targeting rules deciphered by Mohr and colleagues and a modification of the *E. coli* assay originally devised by Matsuura, members of our lab and the Lambowitz lab refined the sequence requirements for retargeting the Ll.LtrB to new sites (figure 2.3) (Matsuura et al., 1997; Guo et al., 2000; Mohr et al., 2000). The Ll.LtrB intron was then retargeted using these refined sequence rules (figure 2.4) and shown to successfully insert into therapeutically relevant targets in mammalian cells (Guo et al., 2000).

Research groups have since applied these rules to generate Ll.LtrB intron libraries in which targeting sequences are randomized yet still satisfy the criteria for mobility. Work by Zhong and Yao together illustrated that retargeted introns (“targetrons”) could efficiently disrupt genes in *E.coli* and that, after mobility, these randomly generated targetrons could then be isolated for use as donor plasmids for their particular insertion site (Zhong & Lambowitz, 2003; Yao et al., 2005).

In 2004, Perutka and colleagues described a computer model that enabled one to search for ideal retrohoming sites within any gene by applying established rules for targeting Ll.LtrB (Perutka et al., 2004). This algorithm has since been employed for rational design of targetrons to targets in *Staphylococcus aureus* and other bacteria (Yao et al., 2006; Yao & Lambowitz, 2007).
Figure 2.3 Sequence requirements for retargeting the L1.LtrB to novel sites.
A. Image from Guo et al, Science 2000 (complete citation in References section). The L1.LtrB intron binding to its natural DNA target sequence. The stars indicate nucleotides important for IEP binding to the target. The horizontal boxes denote nucleotides important for intron RNA binding to the target. The four IEP-related target bases are somewhat fixed whereas the bases used by the intron RNA are flexible so long as the intron’s EBS2, EBS1 and δ are changed accordingly. The arrows indicate cleavage sites by the RNP.
B. The malleability in DNA target site selection. The sense strand of a hypothetical DNA target is shown using the sequence requirements elucidated by Guo and others (Guo et al, Science 2000 with complete citation in References section). One nucleotide in IBS2 and the +5 nucleotide are invariant. The arrow denotes the cleavage site created by the intron RNA during reverse splicing.
Figure 2.4 Ll.LtrB retargeted to bind a therapeutically relevant DNA target.
Image from Guo et al, Science 2000 (complete citation in References section). An HIV-1 DNA target was scanned for potential mobility sites and one result is shown here (sense strand only) with the boxes indicating nucleotides identical to Ll.LtrB’s natural DNA target site. The Ll.LtrB intron was then modified to provide complementary sequence to this selected DNA target site and mobility efficiency was assessed. The vertical line denotes the location of cleavage by the intron RNA.

J.P. Jones, a previous graduate student in our lab, devised a lacZ blue/white screen for Ll.LtrB mobility events in E. coli (Jones et al., 2005). We used this tool to examine the potential utility of retrohoming in gene therapy applications and focused primarily on the specificity of the reaction. We also began experiments for assessing the feasibility of using the Ll.LtrB as a therapeutic in mammalian cells.

2.2 Testing the Specificity of Retrohoming in a Bacterial System

2.2.1 Results

We first used the lacZ blue/white mobility assay with the wildtype Ll.LtrB DNA target sequence to ensure that the assay worked in our hands. We generated a schematic of the experiment and it is shown in figure 2.5. All details are described in the figure legend and in the Materials section, but the general approach can be summarized as RNP production, retrohoming, isolation of the mobilized and immobilized target plasmids via ampicillin-based selection, and quantification of mobility events by blue/white screening.
of the ampicillin-resistant colonies with all steps occurring in one of two different *E. coli* strains (figure 2.5).

Our mobility results using the wildtype Ll.LtrB intron and target were similar to those previously obtained by J.P. Jones and they are provided in table 2.1 as a mean over three different experiments. We also found the IPTG inducible system to be quite leaky as our uninduced sample produced very similar results to those obtained when IPTG was used. However, since only the bacteria carrying the ampicillin resistant target plasmid would have survived to be scored in the blue/white assay and also since the full-length *lacZ* α-complement gene was in the reverse orientation to its promoter in the donor plasmid, the leakiness of the IPTG system was not a major concern.
Two plasmids each with a unique drug resistance gene are used in this assay. The first, “donor”, carries the L.l.ltrB group II intron with wildtype EBS2, EBS1, and δ sequences along with a full-length lacZ α-complement gene in reverse orientation (to prevent expression in the absence of retrohoming) in its domain IV. The IEP (ltrA) ORF has been relocated to a position downstream of the intron in the donor plasmid. L.l.LtrB and IEP expression is controlled by a T7 promoter that contains a lac operator and hence it is regulated by the lac repressor. The second plasmid is called the “target” and contains the DNA target sequences recognized by the wildtype L.l.LtrB group II intron in the reverse orientation relative to the promoter it carries. Downstream of this wildtype target sequence is a deleted version of the lacZ α-complement gene. Both plasmids are co-transformed into HMS174(DE3) E.coli (Novagen) which provides the lac repressor to prevent L.l.LtrB transcription and dual transformants are selected by using both ampicillin and chloramphenicol simultaneously. When IPTG is added, T7 RNA polymerase is expressed in the bacteria and the repressor is also removed from the T7 promoter, which together allow for L.l.LtrB expression and RNP production. The RNPs then mobilize into the target site in the target plasmid and this now places the full length lacZ α-complement gene in the correct orientation for transcription. DNA is isolated from the bacteria and re-transformed into Top10 E. coli (Invitrogen) that can be used for blue/white screening and the transformants are selected for using ampicillin. When x-gal is added, only those colonies with the repaired target plasmid will produce the full length lacZ α-complement necessary to join the host cell-provided lacZ ω-complement and result in blue colonies. Mobility efficiency is then scored at this step using the ratio: (number blue colonies) / (number blue plus white colonies).
Transform into HMS174(DE3) E. coli then initiate T7 transcription by adding IPTG

(Figure continues on next page)
Transform into Top10 E. coli for blue/white screening

Target Plasmid

E. coli

Target Plasmid

Top10 provides

lacz alpha

lacz omega

lacz protein

x-gal

Not Repaired

Repaired
Table 2.1 Repair efficiency of the Ll.LtrB group II intron using an *E. coli* assay

Co-transformations using the above donor and target plasmids were performed according to the experiment outlined in figure 2.5 and mobility assayed by blue/white screening. The backbone plasmids for the donor and target were pACD (derived from pBR322) and pUC19 (with *lacZ* α-complement deletion) respectively. pIntronLacDead is an inactive donor control that was made by deleting portions of the Ll.LtrB catalytic core in domain V. The pIntronLac alone control transformation produced no colonies in the presence of ampicillin. The percent mobility was calculated by counting the colonies on the ampicillin plates at the last step of the protocol (figure 2.5) and each result given here is a mean of three different experiments. Percent mobility = \[\frac{\text{blue colonies}}{\text{blue plus white colonies}}\] x 100.

<table>
<thead>
<tr>
<th>Donor Plasmid</th>
<th>Target Plasmid</th>
<th>IPTG?</th>
<th>Percent Mobility</th>
</tr>
</thead>
<tbody>
<tr>
<td>pIntronLac</td>
<td>pLacTarget</td>
<td>yes</td>
<td>55%</td>
</tr>
<tr>
<td>pIntronLac</td>
<td>pLacTarget</td>
<td>no</td>
<td>52%</td>
</tr>
<tr>
<td>None</td>
<td>pLacTarget</td>
<td>yes</td>
<td>None detected</td>
</tr>
<tr>
<td>pIntronLacDead (catalytic domain deletion)</td>
<td>pLacTarget</td>
<td>yes</td>
<td>None detected</td>
</tr>
<tr>
<td>pIntronLacDead</td>
<td>pLacTarget</td>
<td>no</td>
<td>None detected</td>
</tr>
</tbody>
</table>
We next wanted to examine the specificity of group II-mediated gene repair since off-target integration is a concern for DNA based therapeutic approaches. We sought to determine if the intron designed to repair mutant *lacZ* genes could distinguish between its intended target sequence and the most similar sequences found in the human and *E. coli* genomes. Therefore we searched the human and *E. coli* genome databases for those sequences with the greatest homology to the 17 basepair RNA binding site present in the *lacZ* target sequence. In the human genome, one sequence was found that exactly matched this 17 nucleotide long sequence while in *E. coli* the closest sequence only contained 13 out of 17 base pairs (table 2.2, mutants 1 and 2 respectively). Both of these sequences had alterations in the regions flanking the 17 nucleotide long recognition site that are thought to be important for protein recognition at the insertion site (Guo et al., 2000; Karberg et al., 2001). The 17 basepair RNA binding sequence was found in human chromosome 16 and we used the 45 basepair region encompassing this sequence to construct specificity mutant 1 (table 2.2, mutant 1). Similarly we constructed specificity mutant 2 using the region encompassing the most homologous sequence from *E. coli* (table 2.2, mutant 2). Specificity mutant 3 contains only the 17 nucleotide long sequence from the closest *E. coli* sequence match along with the *lacZ* target sequence derived flanking regions (table 2.2, mutant 3).
Table 2.2 Target mutants constructed from various genomic sequences then used with pIntronLac in a lacZ E. coli mobility assay.

<table>
<thead>
<tr>
<th>Target</th>
<th>Sequence (5' to 3')</th>
<th>E. coli Colony Data</th>
<th>Percent Mobility</th>
</tr>
</thead>
<tbody>
<tr>
<td>LacTarget</td>
<td>acceactegatcgta/ACACATCCATAACCATA/tcatatattat</td>
<td>465 white, 667 blue</td>
<td>49% +/-19%</td>
</tr>
<tr>
<td>Mutant 1</td>
<td>caga-tcatatgt---c/-------------------/atgcaa-g-c</td>
<td>6952 white, 0 blue</td>
<td>None detected</td>
</tr>
<tr>
<td>Mutant 2</td>
<td>tgttga-ttcca---g/-------------------ACCG/ttcccgtatt</td>
<td>4622 white, 0 blue</td>
<td>None detected</td>
</tr>
<tr>
<td>Mutant 3</td>
<td>-------------------/--------------------ACCG/------------------</td>
<td>1409 white, 0 blue</td>
<td>None detected</td>
</tr>
</tbody>
</table>

Only the sense strand is shown here. LacTarget is the same plasmid as pLacTarget used in figure 2.6. It contains the wildtype L1.LtrB 45 bp sequence comprised of the ltrA binding site (lowercase) and RNA binding site (uppercase). Mutant 1 - Searching the human genome uncovered this sequence that contains the LacTarget RNA binding site. Dashed lines denote bases that are identical to LacTarget sequence. Mutant 2 - Searching the E.coli genome uncovered this sequence. It contains sequence that is similar, though not identical, to the LacTarget RNA binding site. Mutant 3 - We changed the E. coli sequence flanking the “RNA binding site” to exactly match the LacTarget ltrA binding sequence. Colony counts represent the sum of 3 independent experiments. Percent mobility = [blue colonies/ (white + blue colonies)] [100].
We then evaluated the ability of the \textit{lacZ} targeted group II intron to distinguish between the genome-derived specificity mutant targets and the intended target sequence in the mutant \textit{lacZ} gene using the group II mobility assay described in figure 2.5. As shown in table 2.2, the intron efficiently mobilized into the intended \textit{lacZ} target DNA, but we could not detect any mobilization into any of the mutant target sites despite the fact that we screened thousands of colonies for such insertions. (A catalytically inactive version of the ribozyme did not mobilize into any target tested.) These results demonstrate that group II RNPs can distinguish between an intended target sequence and the most homologous sequence found in the human and \textit{E. coli} genomes by more than a 1000-fold.

Our lab had also generated a L1.LtrB intron targeted to the first intron of the \textit{\beta}-globin gene as an approach to repair a downstream mutation (IVS2-654) involved in a type of \textit{\beta}-thalassemia (Jones et al., 2005). This intron has been shown to mobilize into the mutant \textit{\beta}-globin target using an \textit{E.coli} assay with subsequent expression of the repaired \textit{\beta}-globin plasmid in mammalian cells confirming the restored product (Jones et al., 2005). We searched the human genome to find a plausible non-specific target for our \textit{\beta}-globin targeted group II intron and found a match on chromosome 14. The sequence contained a region nearly identical to that used in our “wildtype” \textit{\beta}-globin DNA target’s RNA binding site (table 2.3) and we cloned it and the surrounding genome sequence into the target backbone plasmid. Our future plan is to test this \textit{\beta}-globin target mutant in the
*E.coli* tetracycline assay (Jones et al., 2005) to see if our β-globin group II intron will non-specifically integrate into it.

<table>
<thead>
<tr>
<th>Target</th>
<th>Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-globin Wildtype</td>
<td>TGGGTTTCTGATAGGCA <strong>CTGACTCTCTCTGCTTA</strong> TTGGTCTATTT</td>
</tr>
<tr>
<td>B-globin Mutant</td>
<td>---C-GGTATGCCCTGT -------------------C ACCTCAG-GCA</td>
</tr>
</tbody>
</table>

Table 2.3  The wildtype and mutant β-globin DNA targets.  
Only the sense strand is shown. Nucleotides identical to the β-globin wildtype sequence are indicated by dashes. β-globin Wildtype: Human β-globin sequence used previously as a target for a modified GII intron. The 45 bp target sequence is shown with the ltrA binding site in black and the RNA binding site in blue. β-globin Mutant: A nearly identical match to the β-globin target RNA binding site was found on human chromosome 14. We used this and its surrounding sequence to create the 45 bp β-globin mutant target.
2.3 Moving the L. lactis Group II Intron into a Mammalian System

2.3.1 Results

In order for the L.ltrB group II intron to progress as a potential therapeutic agent, its activity needed to be demonstrated in a mammalian system. We chose HEK (human embryonic kidney) 293 and 293T mammalian cells for our studies. Work by others in our lab with the *Tetrahymena thermophila* group I ribozyme had previously shown that T7-based expression in the mammalian cell cytoplasm provided robust levels of ribozyme activity and since we were initially interested in detecting only the group II forward splicing reaction, we also chose this approach (Byun et al., 2003).

A group II intron under the control of a T7 promoter was transfected along with CMV-driven hLtrA and T7 RNA polymerase expression plasmids into 293 cells (figure 2.6). We anticipated that the group II intron would be transcribed in the cytoplasm following nuclear T7 RNA polymerase production and that the hLtrA protein would then bind this pre-splicing cytoplasmic transcript. If forward splicing occurred, we would then attempt to detect it via the non-RNP product of the splicing event. We chose not to use the RNP’s lariat RNA as our forward splicing indicator since we were initially employing an RNase protection assay (RPA) for detection. We felt that the RPA probe might encounter difficulty when trying to bind the structured lariat RNA and this possibility shifted our focus to the spliced exon product instead.
Figure 2.6 Assay for detecting L1.LtrB intron forward splicing in mammalian cells.
We co-transfected three expression plasmids with the group II intron expressed via a T7 promoter and the other two via CMV (the word “pro” denotes promoter). After CMV expression of the T7 RNA polymerase, the group II containing transcript is produced in the cytoplasm. The hLtrA (the bacterial ltrA ORF now with eukaryotic codon usage) protein then binds this transcript and forward splicing can proceed. If forward splicing occurs, two products are made: the RNP and the spliced exons. We selected the spliced exons as our indicator for forward splicing events in the RPA experiments since the other product’s lariat structure might have impeded probe binding in the assay.
We designed two probes for the RPA assay, one called “react” that we used for detecting the spliced exons (and hence the group II forward splicing event) and the other termed “unreact” for unspliced transcript detection. The various RPA protected fragments that could result from using either of these two probes with or without forward splicing events are shown in figure 2.7.

![Figure 2.7 Resulting protected RNA fragments (probe binding sites and gel schematics) after using either the “react” or “unreact” RPA probes in the presence/absence of Ll.LtrB forward splicing in mammalian cells.](image)

The “react” probe (red) detects the forward spliced product. The “unreact” probe (light green) detects the unspliced T7 transcript. All numbers denote the length of each fragment (in nucleotides for the RPA probe binding schematic and in basepairs for the dsRNA protected fragments shown on the gel). The black 50 nucleotide fragment is used to distinguish full length probe from digested probe. The spliced product is a solid light blue line and the unspliced product contains an intervening light blue dashed line to indicate the presence of the group II RNA. The hypothetical gel includes the sizes of the undigested full length probes for completeness although these would be seen in different lanes in actual experiments.
We first tested our probes and RPA assay by using \textit{in vitro} transcribed forward spliced and unspliced RNAs (figure 2.8). The radiolabeled probes were co-precipitated and hybridized to the positive control RNAs and subjected to RNase treatment. We detected the expected sized products upon denaturing polyacrylamide gel electrophoresis for both cases and maintained that probe’s signal in the combined presence of both \textit{in vitro} transcribed RNAs (figure 2.8, bands 1 and 3). The undigested full-length probes were also the anticipated sizes (figure 2.8, bands 2 and 4) and neither probe bound non-specifically to yeast RNA (figure 2.8 compare lanes 6/7 and lanes 14/15).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2_8.png}
\caption{Testing the react and unreact controls for the RPA experiment.}
\end{figure}

Each radiolabeled probe was first co-precipitated and hybridized to \textit{in vitro} transcribed forward spliced and unspliced RNA controls. These samples were then subjected to RNase treatment and the resulting bands on the gel are labeled 1 and 3 respectively (lanes +R and +U). The lanes R/+U and +U/+R represent the specific probe hybridized to a 1:1 mixture of each \textit{in vitro} transcribed splicing control. This was used to crudely assess the sensitivity of the probe. The F lanes are undigested full length probe with bands 2 and 4 on the gel representing the react and unreact probes respectively. The Y lanes are non-specific yeast RNA co-precipitated with the specific probe and either undigested (Y1) or subjected to RNase treatment (Y2). This was to ensure that our probes were specific to their intended targets. The L lanes are radiolabeled RNA markers with the band sizes written alongside one example. The same ladder was used for each marker lane. The B lanes are blank lanes.
With these controls established, we co-transfected our plasmids as illustrated in figure 2.6 and harvested cytoplasmic RNA 48 hours later. We used a total of 5ug RNA for each RPA sample and an example using different amounts of react probe is shown in figure 2.9. Using 2 fmoles of the react probe with the \textit{in vitro} transcribed react RNA positive control produced the expected 213bp protected fragment (figure 2.9, lane 2). The unspliced control RNA produced the expected 58bp protected fragment, but the 155bp fragment from this RNA was not detected using the react probe (figure 2.9, lane 3). The undigested full length react probe (figure 2.9, lanes 4 and 10) had the expected 263nt fragment and two negative controls were also as expected (no bands seen when either the group II intron or the T7 polymerase expression plasmids were absent – figure 2.9, lanes 5, 6, 11, and 12).

However, the mix control in lanes 8 and 15 produced what appeared to be the 213bp protected fragment. This was unexpected because the mix control was generated by combining pre-lysis cells previously separately transfected with either hLtrA or group II plasmids (plus the T7 polymerase expression plasmid in both cases). The 213bp fragment seen in the mix controls either signifies a false positive result, perhaps by the react probe binding an unspliced transcript with the inner portion (the dashed light blue line in figure 2.7) looping out, or the group II intron RNA truly did forward splice without the LtrA protein, perhaps by either using a substitute protein while in the 293 cells or else during a high-salt step during RNA workup. These possibilities will be explored further in section 2.4.
Figure 2.9 Using different amounts of react probe to detect LI.LtrB forward splicing in mammalian cells.

Cytoplasmic RNA from transfected 293 cells was co-precipitated and hybridized to different amounts of radiolabeled react probe (amounts given in fmoles and noted above the relevant lanes) and used in an RPA. The * refers to the 12E lane where 12 fmoles probe were used. +R and +U refer to the in vitro transcribed forward spliced and unspliced control RNAs respectively. F is the undigested full length react probe, -GII is a co-transfection without the LI.LtrB group II intron plasmid included, and -T7 is a co-transfection without the T7 expression plasmid included. Mix is another control (described in the text). The experimental lanes, E, are co-transfections where all three plasmids (GII, T7, and hLtrA) were used. L is the radiolabeled RNA molecular weight marker and the band sizes are indicated by the arrows.
The experimental lanes (figure 2.9, lanes 7, 13, and 14) originated from the same transfection and included all three plasmids. Lane 7 represents hybridization with 2 fmoles probe, lane 13 is 12 fmoles, and lane 14 is 6 fmoles. All experimental lanes produced the 213bp protected fragment and would seem to suggest that forward splicing occurred in the mammalian cells, however the mix control result makes this finding inconclusive. The unreact probe at 2 fmoles produced the expected band sizes (data not shown) for the undigested full length (354nt), *in vitro* transcribed unspliced RNA (304bp), and the following co-transfection controls: no hLtrA plasmid (304bp), no group II plasmid (304bp), and the mix control that was prepared as described previously (304bp).

We then attempted to detect the forward spliced product via RT-PCR rather than RPA and used a 293 cell line that was stably transfected with the hLtrA ORF (kindly provided by Alan Lambowitz). We co-transfected our group II and T7 polymerase expression plasmids into these 293hLtrA cells and harvested total RNA 48 hours later. We then used 3 different primer sets in RT-PCR reactions with each set designed to distinguish between forward spliced and unspliced RNA transcripts. Our initial attempts to detect forward splicing in the cells were unsuccessful so we decided to examine the functionality of the intra-cellularly expressed transcript and also assess levels of hLtrA protein production by the cell line.

For the first aspect, we harvested total RNA from 293hLtrA cells that had been co-transfected with group II and T7 polymerase expression plasmids 48 hours earlier. We
used a portion of this total RNA in a high-salt self-splicing assay that allows the intron RNA to forward splice in vitro in the absence of the ltrA protein (Saldanha et al., 1999). The monovalent and divalent cations allow the RNA to fold into a catalytically active structure and self-splice after incubation at 37°C. We then used ethanol precipitation to remove the salts and concentrate the RNA. This RNA, along with aliquots of total RNA that were not subjected to high salt treatment, was used in the RT-PCR reactions and the products were run on an agarose gel (figure 2.10). We found that the group II transcripts produced within the mammalian cells were able to forward splice under in vitro self-splicing conditions and hence they were likely functionally viable substrates for protein assisted forward splicing in vivo (figure 2.10).
Figure 2.10  The group II Ll.LtrB transcript produced in mammalian cells will forward splice *in vitro* under high salt conditions. The group II and T7 polymerase expression plasmids were co-transfected into 293 cells that had a stably integrated copy of the hLtrA ORF. Total RNA was harvested 48 hours later and the RNA was incubated under high-salt conditions for 1.5 hours. The RNA was then ethanol precipitated and used in a two-step RT-PCR reaction. Three different primer sets were used to verify the presence of both unspliced and forward spliced transcripts. The “U” lanes are total RNA from the co-transfections that was not treated with the high-salt conditions. All boxed bands in the gel were confirmed by sequencing. The unspliced and forward spliced RT-PCR anticipated product sizes for each primer set were as follows: Green unspliced - approx. 1.5kb, green forward spliced - 610bp. Yellow unspliced - approx. 1.5kb, yellow forward spliced - 547bp. Blue unspliced - approx. 1.4kb, blue forward spliced - 424bp. “L” denotes molecular weight marker lane.
Since we had been unable to detect the forward splicing event *in vivo*, we next assessed the levels of hLtrA protein being made by the 293hLtrA cell line. We also constructed two new expression plasmids that placed the intron RNA and hLtrA ORF in *cis* (figure 2.11). Work by Saldanha and colleagues had shown that the bacterial LtrA protein was unstable when unbound to the intron RNA and we hypothesized that this perhaps contributed to our inability to detect forward splicing in mammalian cells (Saldanha et al., 1999). Since the 293hLtrA stable cell line (and all prior plasmid transfections) expressed the hLtrA separately from the intron RNA, the translated cytoplasmic protein was unstable until it encountered the T7 expressed intron RNA. Perhaps if we transcribed both the intron RNA and the hLtrA RNA from the same template (using either a T7 promoter plus an IRES for cap-independent translation or by converting all to an RNA polymerase II system), this would place the two components in closer proximity and encourage binding. The protein would thus be stabilized and forward splicing might be more apt to occur.
Figure 2.11  Providing the hLtrA protein in cis.
Previous work (Saldanha et al, Biochemistry 1999. Complete citation in References) had shown that the LtrA protein was more stable when bound to the intron RNA. We therefore designed two plasmids, each expressing both the RNA and protein ORF as one transcript, to put the components in closer proximity and hopefully stimulate protein binding to the intron RNA. pGII_IREShLtrA was expressed in the cytoplasm via a T7 promoter and the hLtrA protein was generated by IRES-directed translation. The other plasmid, phLtrAGII, used RNA polymerase II for nuclear transcript expression and the hLtrA polyA signal was moved downstream of the intron sequence to ensure production of a single transcript. This capped/polyadenylated transcript was then shuttled to the cytoplasm for ribosomal translation of the hLtrA (using a stop codon to prevent translation of the downstream intron). As with all group II intron plasmids used in the mammalian studies, both plasmids contained domain IV deletions that removed the natural LtrA ORF.
Western blotting with an antibody to the LtrA protein and using an *in vitro* reconstituted group II RNP as a positive control (both reagents courtesy of Alan Lambowitz) showed that although the protein was present in both the 293hLtrA cells and a 293T co-transfection that included pGII_IREShLtrA, its levels were quite low (figure 2.12). The *cis* IRES construct did not appear to increase hLtrA protein levels suggesting that either the *cis* approach did not stimulate increased protein/RNA binding or perhaps that the IRES-based translation was ineffective. We decided to continue our assay for forward splicing using either *cis* construct (pGII_IREShLtrA or phLtrAGII) in 293T transfection experiments followed by RT-PCR with the primers described in figure 2.10. The products were assessed via agarose gel electrophoresis and no forward spliced product was detected.

![Figure 2.12 Western blot illustrating low levels of hLtrA protein production in mammalian cells.](image)

Protein extracts from either 293T cells co-transfected with pGII_IREShLtrA (see figure 2.11) and T7 polymerase plasmids or from 293 cells stably transfected with an hLtrA ORF were isolated and used in a western blot experiment. An antibody to the LtrA protein was used and the positive control was an *in vitro* reconstituted L1.LtrB RNP.
2.4 Conclusions from Sections 2.2 and 2.3

We have shown that the Ll.LtrB group II intron mobilizes specifically into DNA targets using an *E. coli* assay to detect retrohoming. Even when presented with plausible and in some cases highly similar DNA target sites, Ll.LtrB did not mobilize into these non-specific sites (table 2.2). We are now preparing to use a tetracycline-based *E. coli* assay to determine if an Ll.LtrB intron retargeted to a β-globin DNA site will show the same high degree of specificity during retrohoming.

The ability to select and precisely integrate into a given DNA target is a valuable trait for potential genetic therapies. Targeted integration into the genome allows for permanent addition, disruption, or even repair of a gene of interest. However, many DNA-based therapeutic approaches utilize integrating viral vectors that can lead to deleterious outcomes when the transgene disrupts a vital gene or activates a potential proto-oncogene (McCormack & Rabbitts, 2004). The Ll.LtrB inserts site-specifically and now has been shown to maintain this specificity even when faced with nearly identical target sites in an *E. coli* assay (table 2.2, mutant 3).

The next challenge was transitioning the Ll.LtrB activity into mammalian cells. Our lab, in collaboration with Alan Lambowitz’s group, has been able to transf ect or microinject reconstituted RNP particles into mammalian cells and detect retrohoming into therapeutically relevant plasmid DNA targets (Guo et al., 2000). We now attempted to express the Ll.LtrB components within mammalian cells and allow the RNP to form *in*
We began by assaying for the first step, the forward splicing reaction, using an RNase protection experiment. Although our probes for detecting the unspliced and forward spliced transcripts appeared to be working (figure 2.8), we unexpectedly discovered the 213bp protected fragment in our mix control (figure 2.9).

The mix sample contained hLtrA and group II intron (ltrA ORF deleted) transcripts from separate transfections that were mixed immediately prior to cell lysis and RNA isolation. Since any hLtrA protein produced intra-cellularly was likely denatured during plasma membrane lysis or the RNA isolation steps, we did not expect to see the 213bp “forward spliced” protected fragment in the mix control. However, we did and the band was of comparable intensity to the 213bp bands seen in the experimental lanes (figure 2.9).

The 213bp fragment in the mix control may be a false positive result. Perhaps the react probe was binding an unspliced transcript with its inner section (the dashed light blue line in figure 2.7) looped out. This mechanism for false positives in RNase protection assays has been suggested by others and is very possibly the cause of the 213bp bands seen in mix and experimental lanes (Frantz et al., 1999; Takahara et al., 2002).

An alternative possibility is that the group II intron RNA truly did forward splice without the LtrA protein, perhaps by using a substitute protein while in the mammalian cells. Mohr and colleagues recently showed that certain host cell proteins in yeast and other organisms may help RNA fold into its catalytically active state and then, once the
RNA is folded, are no longer necessary for forward or reverse splicing (Mohr et al., 2006). It is thought that these proteins act by preventing the folding RNA from stalling at any kinetically-favorable conformation until the final active structure is achieved (Herschlag et al., 1994). Although the specific protein they were studying, CYT-19 (found in *N. crassa* and with a homolog present in *S. cerevisiae*), did not appear to act as a chaperone for Ll.LtrB RNA, this does not discount the possibility that a mammalian cell protein might be able to fulfill this function. Perhaps a protein in our human cell line is assisting the intron RNA in its folding to a catalytically active state without the hLtrA present. However, our RT-PCR experiments should have then detected these forward spliced transcripts and since they did not, this makes this possibility less likely.

The RPA mix control forward splicing result may also be real if our unspliced transcript was able to self-splice during any of the post cell-lysis processing steps. As was described earlier, high-salt treatment stabilizes the intron RNA’s catalytically active state and leads to self-splicing. Perhaps one of the steps during the RPA experiment produced conditions favorable for self-splicing whereas no similar conditions were attained during the RT-PCR experiments. This would then lead to a true positive result in the RPA samples containing the intron RNA (including the mix control) and a true negative result in the RT-PCR.

**2.5 Materials and Methods for Sections 2.2 and 2.3**

**Donor and Target Plasmid Construction for *E.coli* Mobility Assay**
The pLacTarget plasmid was generated by deleting the 3’ portion of the α-complement of the lacZ gene in pUC19 between the EcoRI and NarI restriction sites and inserting the natural intron target site immediately upstream of this deletion in the antisense orientation to the promoter. Each target mutant plasmid was made by swapping the natural intron target site with one of the three mutants described in the text. The pIntronLac plasmid was generated by inserting the open reading frame (ORF) for the α-complement of the lacZ gene in the antisense orientation into the MluI site in domain IV of the L. lactis group II intron present on the pACD WtΔ-ORF+ORF plasmid (Guo et al., 2000). This plasmid is derived from pBR322 and thus has a different origin of replication from pUC19. The inactive version of pIntronLac was generated by deleting sequences in domains V and VI of the intron between the PstI and SnaBI restriction sites.

**Gene Repair Analysis in the E. coli Mobility Assay**

For lacZ gene repair studies, E. coli HMS174 (DE3) (Novagen) was cotransformed with the lacZ target plasmid and the intron plasmid and incubated in LB with ampicillin (50 ug/ml) and chloramphenicol (25 ug/ml) overnight at 37 degrees Celsius (37°C) with shaking. The next morning, 5 ml LB media with 50 ug/ml ampicillin and 25 ug/ml chloramphenicol was inoculated with 100 ul of the overnight culture and shaken at 37°C until the OD$_{600}$ was approximately 0.2. Group II intron expression was then induced by addition of IPTG (100 uM) and the cultures incubated with shaking at 37°C for 1 hour. Plasmid DNA was then isolated from the cells using a QiaQuick miniprep kit (Qiagen). LacZ gene repair was assessed by transforming E. coli TOP10
cells with the miniprep plasmid DNA and scoring for ampicillin resistant colonies that were blue when grown on bacterial plates containing IPTG and X-gal.

**Plasmids Used in Mammalian Cell Experiments**

All Ll.LtrB group II introns used in these experiments had a domain IV deletion to remove the endogenous ltrA ORF. The T7 RNA polymerase expression plasmid contained both CMV and T7 promoters upstream of the T7 RNA polymerase ORF. phLtrA was a CMV-driven expression plasmid that contained the LtrA ORF with sequence alterations for human codon usage during translation in the mammalian cells. The Ll.LtrB group II expressing plasmids were one of three varieties and all based on the pcDNA3.1+hygro backbone plasmid (Invitrogen). The first was the T7 promoter driven group II plasmid and the backbone CMV promoter was deleted so only T7-based expression would occur. The other two plasmids are described in the text (and in figure 2.11) and both were designed to produce the intron RNA and hLtrA RNA as one transcript.

**Transfections**

Human embryonic kidney (HEK) cells were plated onto either 6-well plates or 100mm dishes and allowed to grow overnight to achieve 40-80% confluency. The media (DMEM (Invitrogen), 10% fetal bovine serum (HyClone), 1x penicillin/streptomycin (Invitrogen)) was replaced the next morning immediately before transfection. About 5ug (6 well plates) or 10-20ug (100mm dishes) of total DNA was suspended in 200ul Opti-MemI (Invitrogen) and allowed to interact with Superfect lipid transfection reagent
(Qiagen) for 10 minutes at room temperature. The DNA to Superfect ratio was 2:1 for all transfections. The DNA/Superfect complexes were then dripped onto the cells and placed in a humidified $37^\circ C$ incubator with 5% CO$_2$ for 48 hours.

**RNA Isolation**

48 hours after transfection, cells were trypsinized and either cytoplasmic or total RNA was harvested using Qiagen’s RNeasy Miniprep kit. This protocol included a DNase treatment step. The mix controls were made after trypsinization, but before cell lysis and contained 1/2 the cells from each relevant well. The total RNAs were quantified using a NanoDrop spectrophotometer (NanoDrop Technologies).

**High-Salt Treatment**

We used the methods described by Saldanha (Saldanha et al., 1999). Briefly, 1-2 ug total RNA was added to a high salt buffer containing the following reagents in their final concentrations: NH$_4$Cl (1.5M), MgCl$_2$ (100mM), TrisHCl pH7.5 (40mM) and DTT (10mM). The samples were heated to 65°C for 1 minute then placed in a 37°C water bath for 1.5 hours. After incubation, the RNA samples were cleaned and concentrated via ethanol precipitation.

**RT-PCR**

Equal amounts of RNA (cytoplasmic or total, high-salt treated or untreated - depending on the particular experiment) were used in two-step RT-PCR reactions. We used Invitrogen’s Superscript III reverse transcriptase and an RNase inhibitor (Invitrogen) and incubated the reverse transcription reactions at 55°C for 1 hour. We then
used 1/10 of each RT reaction in a PCR reaction using Platinum taq (Invitrogen). The annealing temperatures varied according to the primer sets used but the general set-up was as follows: 95°C for 5 minutes, 25 cycles of (95°C 45 seconds/55-59°C 30 seconds/72°C 30 seconds) then 7 minutes at 72°C with a final hold at 4°C. The products were then visualized using ethidium bromide staining and agarose gel electrophoresis. If the product was sequenced (noted in the Results section), Qiagen’s QiaQuick Gel Extraction kit was used for DNA isolation.

**RNase Protection Assay**

The Ambion Company’s RPAIII kit was used for all experiments.

Probe and RNA Co-Precipitation/Hybridization: Identical amounts of total RNA (in ug) were diluted in 80ul of water per tube. The amount of total RNA used in each experiment varied, but an average was about 2.5ug. A number of controls were used during the study: water-alone, yeast RNA without/with RNase digestion, and *in vitro* transcribed forward spliced or unspliced RNA. The purpose of the controls was to assess the integrity of the full-length probe, ensure the probe’s specificity for its target, and determine the extent of RNase digestion. Each RNA sample was co-precipitated with either 2, 6 or 12 fmol (amount is specified in relevant portion of Results) radiolabeled (α-32P-CTP) polyacrylamide gel purified full-length RNA probe using 0.5M ammonium acetate and 100% ethanol. The tubes were placed in an -80°C freezer for 20 minutes then the RNA was pelleted using a 4°C microcentrifuge set at 13,000 RPM for 20 minutes. The supernatant was removed and the pellet allowed to air-dry for 5 minutes. The co-
precipitated probe and RNA sample was resuspended in 10-20ul of the hybridization buffer provided in the RPA III kit and allowed to hybridize overnight at either 42°C or 56°C.

RNase Digestion of Unbound RNA: The RNA that did not hybridize to the probe was digested using a combination of RNase A and RNase T1. A 1:150 ratio of enzyme to digestion buffer (provided in the RPAIII kit) was used. The digestion was allowed to proceed for 30 minutes at 37°C and stopped using the RPAIII kit’s RNase Inactivation and Precipitation solution. The samples were ethanol precipitated, resuspended in 6ul Gel Loading Buffer II (provided in the RPAIII kit), and loaded onto a 6% denaturing polyacrylamide gel along with a γ-32P-ATP labeled RNA molecular weight marker (Ambion). The gel was then dried using vacuum and heat (80°C) for 2 hours and exposed overnight on a phosphor screen. The screen image was then captured using a Storm Phosphorimager (Molecular Dynamics/GE Healthcare) and the results were analyzed using ImageQuant software (Molecular Dynamics/GE Healthcare).
Chapter 3: RNA Repair Using Spliceosome-Mediated RNA Trans-Splicing (SMaRT) Technology

3.1 Introduction to SMaRT Technology

*Trans*-splicing is a naturally occurring phenomenon in some viruses, prokaryotes, lower eukaryotes, and possibly even in mammalian cells (Ziff & Evans, 1978; Michel & Dujon, 1983; Milhausen et al., 1984; Blumenthal, 1988; Kohchi et al., 1988; de Souza, 1991; Finta & Zaphiropoulos, 2002). The mechanism varies from the self-catalyzed reaction of the discontinuous group II intron to the spliceosome-driven splice leader approach used by nematodes and trypanosomes (Kohchi et al., 1988; de Souza, 1991). *Trans*-splicing can also be artificially derived by either modifying a natural splicing process, as in the group I ribozyme, or by manipulation of the *cis*-splicing machinery (Sullenger & Cech, 1994; Puttaraju et al., 1999). The latter is termed Spliceosomal Mediated RNA *Trans*-Splicing or “SMaRT” and it has been studied as a potential therapeutic for diseases including cystic fibrosis, hemophilia A, a form of epidermolysis bullosa, and tauopathies (Mansfield et al., 2000; Liu et al., 2002; Chao et al., 2003; Dallinger et al., 2003; Rodriguez-Martín et al., 2005). The general methodology has also been used as a possible cancer therapeutic via intra-tumoral reconstitution of a toxin RNA (Nakayama et al., 2005).

SMaRT technology uses a PTM (**Pre**-*Trans*-**Splicing** **Molecule**) to “trick” the spliceosome into using it as a substrate for splicing (Puttaraju et al., 1999). It achieves
this by blocking the splicing signals on the nascent pre-mRNA transcript via base-pairing while concurrently providing the spliceosome with alternative splicing sequences (Figure 3.1). The PTM sequence complementary to the target RNA is called the “binding domain” (B.D.) and it is followed by non-complementary “spacer” sequence. The spacer affords flexibility to the PTM’s splicing region hence increasing the chance the region will be in a favorable configuration for loading of endogenous splicing machinery. The PTM’s splicing sequences may be either 5’ or 3’ splice sites and the cargo sequence is at the end of the molecule. The objective is for the spliceosome to recognize the PTM’s splice sites instead of the endogenous RNA’s signals. Achieving this goal results in the trans-splicing of the PTM’s cargo onto the target RNA and, in therapeutic applications, repair of the mutant endogenous RNA. The technology is desirable in that it preserves the endogenous regulatory milieu and requires delivery of only a portion of the transcript, the latter important for some limited capacity viral vectors.
Figure 3.1 Exon replacement via trans-splicing using Spliceosome-Mediated RNA Trans-Splicing Technology (SMaRT).

A. Mechanism for 3’ trans-splicing RNA repair. A 3’ Pre-Trans-splicing Molecule or “PTM” RNA basepairs to the target RNA and conceals the endogenous 3’ splice site region (branchpoint, polypyrimidine tract, 3’splice site (3’ss) inclusive) while concurrently providing its own 3’ region that, if used by the spliceosome, introduces a therapeutic RNA sequence. B.D. refers to the binding domain; the sequence that basepairs to the target RNA. The spacer sequence provides pliability to the PTM’s splice region in hopes of optimizing its configuration for loading of spliceosomal components. The PTM’s splice region contains the branchpoint (circle), polypyrimidine tract (triangle) and 3’ss. B. Mechanism for 5’ trans-splicing RNA repair. The 5’PTM functions similarly to the 3’PTM with some important differences. The 5’PTM carries 5’ss sequences and introduces the therapeutic sequence upstream of the target binding site hence its use is described as 5’ exon replacement.
A

Target Pre-mRNA

5' Pre-Trans-splicing Molecule (5'PTM)

EXON 1

EXON 2

EXON 3

EXON 4

3' Exon Replacement via Trans-Splicing

EXON 1 EXON 2 THERAPEUTIC

B

Target Pre-mRNA

5' Pre-Trans-splicing Molecule (5'PTM)

EXON 1

EXON 2

EXON 3

EXON 4

5' Exon Replacement via Trans-Splicing

THERAPEUTIC EXON 3 EXON 4
PTMs can be crafted to replace RNA sequence either 5’ or 3’ to their binding sites on the endogenous pre-mRNA targets; these are described as 5’PTMs and 3’PTMs respectively (Puttaraju et al., 1999; Mansfield et al., 2003). 5’PTM usage is termed “upstream exon replacement” whereas the converse downstream process utilizes 3’PTMs. 3’PTMs have been employed to correct a number of different clinically-relevant mutations using delivery routes spanning from in vitro applications to mouse models of disease (Puttaraju et al. 1999; Mansfield et al. 2000; Kikumori et al. 2001; Puttaraju et al. 2001; Liu et al. 2002; Chao et al. 2003; Dallinger et al. 2003; Tahara et al. 2004; Liu et al. 2005; Rodriguez-Martin et al. 2005). A technique nearly identical to SMaRT, termed “STS” for Segmental Trans-Splicing, has also been studied as a possible cancer therapeutic via trans-splicing reconstitution of mRNA toxins in cancer cells (Pergolizzi et al., 2003; Nakayama et al., 2005). Finally, SMaRT and 3’PTMs have been used as tools for molecular imaging within living animals (Bhaumik et al., 2004).

The initial paper describing SMaRT technology was published in 1999 by Puttaraju and colleagues (Puttaraju et al., 1999). They demonstrated the use of 3’PTMs in a number of experiments ranging from in vitro splicing assays to cancer tumor modeling in mice. The 3’PTM designs were first tested in a HeLa nuclear extract in vitro splicing assay using T7 transcribed PTMs and the β-subunit of human chorionic gonadotropin gene 6 as the RNA target. Trans-splicing was detected using RT-PCR amplification followed by sequence confirmation of the products. One of the designs utilized a sequence called a “safety stem” that provided an additional safeguard against non-
specific trans-splicing. The safety version of the binding domain contained some homology to the PTM’s splice regions and this created a stem structure in the PTM RNA. Once the PTM encountered its true RNA target, the safety binding domain sequence would “release” from the structure and preferentially bind the target due to greater sequence homology at that site. They showed that the safety decreased non-specific trans-splicing using an unrelated target in their in vitro splicing assay.

Later work by Puttaraju et al. explored the effect of binding domain length on trans-splicing efficiency and specificity (Puttaraju et al., 2001). The study found that increasing the binding domain led to increased efficiency and specificity. They postulated that the former was due to the longer binding domain’s resiliency in a helicase-rich cellular environment and that specificity improvements were the result of safety-like structures forming in longer RNAs. Interestingly, they also found that regions of non-binding within the binding domain actually improved efficiency of trans-splicing and suggested that perhaps this was because it reduced the potential for ADAR responses (Adenosine Deaminases that Act on RNA).

3’PTMs have been tested in a variety of ways in animal models. Liu and colleagues delivered 3’PTMs via adenovirus to human deltaF508 CFTR xenografts on rodent hosts and detected restoration of CFTR Cl- channel function to 22% of normal (Liu et al., 2002). Tahara et al. transplanted bone marrow previously infected with 3’PTM-carrying lentiviruses into irradiated mice that were producing mutant CD40L transcripts (Tahara et al., 2004). Transplanted mice expressed the trans-spliced RNA
according to RT-PCR experiments and they also showed some restoration of immunoglobulin isotype switching capabilities (CD40L mutants usually have none (Callard et al., 1993; Kroczek et al., 1994)). In addition, CD4+ T-cells from these treated mice began expressing cell surface CD40L and the mice were able to mount some immune response to challenge with *P. carinii*. Chao and colleagues used either direct injection of PTM expression cassettes or adenoviral infection to deliver a 3’PTM to mice that were homozygous mutant for the factor VIII gene (Chao et al., 2003). Factor VIII is involved in the coagulation pathway and its absence leads to hemophilia A. Among other assays, the researchers used a tail-clip challenge to demonstrate restoration of normal factor VIII function after 3’PTM treatment. The challenge was performed 3 weeks post-PTM delivery and they found that 8 of 10 PTM-treated mice survived.

Unlike 3’PTMs, the study landscape for 5’PTMs is quite barren. Codony and colleagues tested a number of different splice donors for *trans*-splicing to c-H-ras pre-mRNA in hopes of modulating alternative splicing (Codony et al., 2001). However, only one of the splice donors used was a targeted molecule (via a 70nt binding domain).

Mansfield and colleagues have published the only in-depth study of targeted 5’PTMs to date (Mansfield et al., 2003). They first used a 5’PTM targeted to intron 10 of a CFTR minigene target carrying the deltaF508 mutation in exon 10. Using RT-PCR and an ion channel functional assay, Mansfield demonstrated that co-transfection of 5’PTM and mutant CFTR minigene target expression plasmids into HEK293 cells led to *trans*-spliced RNA product and some restoration of CFTR function.
The researchers then designed a number of different 5’PTM expression cassettes to test how various aspects of PTM design affected trans-splicing efficiency in mammalian cells. For these experiments, they used a lacZ repair assay to detect trans-splicing and determined restoration of function via β-galactosidase activity. They examined different binding domain positions along the target, PTM spacer lengths, and how incorporation of an intron within the PTM’s cargo influenced trans-splicing activity. Binding domains targeted near the downstream 3’ splice site region proved the most effective and they suggested that perhaps the utility of this design feature stemmed from their targeting of a long intron. They postulated that perhaps because of RNA polymerase II/splice factor interactions, the areas near intron/exon junctions might be “crowded” and a large intron provides targeting sites outside the fray that are still near enough to remain involved in splicing activities.

The incorporation of an intron into the PTM’s cargo also greatly enriched trans-splicing levels. The researchers suggested that this was perhaps because the intron improved the exon definition of the large cargo sequence and hence made it more attractive to the spliceosome. The intron was also thought to provide benefits unrelated to splicing such as improved transport to the cytoplasm and increased translation of the mRNA.

Our study explores the effects of the targets themselves and also uses the 5’PTM in a novel context, that of repairing a natural first exon. Our models for this work are β-globin mutations in the first exon or intron that lead to either sickle cell anemia or β-
thalassemia respectively. We have examined trans-splicing activity using both HEK293T mammalian cell models and primary erythroblast precursors from sickle cell patients.

3.2 Using 5’PTMs to repair the sickle cell mutation in mammalian cell culture experiments

3.2.1 Summary

We have shown that 5’PTMs can be used to correct the first exon of a therapeutically relevant gene in HEK293T mammalian cell co-transfection experiments. Using a genomic mutant copy of the \( \beta \)-globin gene, we found that 5’PTMs can trans-splice to repair the defect causing sickle cell anemia at the RNA level. This is the first time a 5’PTM has been targeted specifically to correct a disease-causing mutation in a gene’s first exon.

3.2.2 Introduction

Mansfield and colleagues provided the original publication regarding 5’PTM design (see section 3.1, (Mansfield et al., 2003)) and we used this work as a guide in developing our 5’PTMs. However, unlike the Mansfield study target, our \( \beta \)-globin target pre-mRNA intron 1 was short and the cargo we were introducing (repaired \( \beta \)-globin exon 1) was also quite small. Furthermore, we were attempting to repair a gene’s natural first exon, which had not been explored previously.
We also tried to consider the potential impact of nuclear dsRNA silencing processes acting on our PTM binding domain/target interaction and designed our binding domain accordingly (by using a fairly short sequence). In a similar vein, we considered the potential for dsRNA formation between our PTM’s binding domain and its cargo sequence and designed one construct in an attempt to minimize this (5’PTMCodon). This construct also allowed for improved distinction between repaired and mutant RNAs as well as provided an opportunity to eliminate any GT dinucleotides that could result in cryptic splicing. Our final 5’PTM designs were introduced into HEK293T (human embryonic kidney) mammalian cells and assayed for their ability to repair βγ-globin transcripts.

3.2.3 Results

Three 5’PTM constructs were designed (Figure 3.2) and cloned into the RNA polymerase II expression vector pEGFPN1 (Clontech). The PTMs varied by 5’ splice site (5’ss) sequence used, spacer length, and inclusion of a potential intronic splice enhancer sequence (“DISE”, Seth, P. and Garcia-Blanco, M. personal communication). All 5’PTMs had a 43 nucleotide (nt) binding domain with the small size used in an attempt to avoid formation of long dsRNA within the nucleus. One construct, 5’PTMCodon, was designed using the degeneracy of the genetic code to make the distinction between cis and trans-spliced products clearer at the nucleotide level. The β-globin exon 1 codons chosen for alteration in 5’PTMCodon originally contained GT dinucleotides (potential
cryptic 5’ss within the PTM) and the alternative codons were selected based on their usage levels in mammalian cells. The alternative codons in 5’PTMCodon also allowed us to minimize potential dsRNA formation between binding domain and cargo sequences. A PTM-specific 5’ tag sequence was added to all constructs to further discriminate between cis and trans splicing events.

Each 5’PTM construct was co-transfected along with genomic β^3^-globin expression vectors into HEK293T cells and trans-splicing was assessed 48 hours later using total RNA in a one-step RT-PCR reaction. Using trans-splicing specific primers followed by sequence confirmation, we detected repair of the sickle β^-globin transcript using each of the three 5’PTMs (Figure 3.3 A). The “mix” control, comprised of PTM and β^-globin transcripts expressed in separate cells then mixed before cell lysis, did not produce the amplification product indicating that the trans-splicing event happened only within the co-transfected cells. Figure 3.3 B shows that although the mix controls did not produce a trans-spliced RT-PCR product, the PTM and β^-globin transcripts were both present and able to be independently amplified.
Figure 3.2. 5’PTM constructs.

All 5’PTMs were generated via RNA polymerase II expression plasmids within HEK 293T cells. A 5’ tag sequence was added to each PTM to improve distinction between cis and trans-spliced products and alternative codons for exon 1 were used in 5’PTMCodon for this same purpose. The codons chosen for alteration in 5’PTMCodon originally contained GT dinucleotides (which could act as cryptic 5’ss in the PTM). The alternate codons introduced at these sites were selected based on their usage levels in mammalian cells. A 43 nucleotide (nt) binding domain was used in all cases and it bound the last 11nt of exon 1 through the first 32nt of intron 1. The three 5’PTMs each had a unique 5’ splice site (5’ss) denoted here as A, B, and C (see Methods for details). 5’PTMDISE was further distinct in that the spacer length used was 12nt rather than 84nt as in the other constructs. This 5’PTM also carried a 45nt putative intronic splicing enhancer element called DISE (Seth, P. & Garcia-Blanco, M. personal communication) downstream of its 5’ss. The symbol * denotes the location of the sickle cell point mutation.
Figure 3.3. Sickle \(\beta\)-globin transcripts are corrected via 5’PTM \textit{trans}-splicing.

A. \(\beta^s\) RNA is repaired via 5’PTM \textit{trans}-splicing in 293T cells. 5’PTM (5’PTM1, 5’PTMDISE, & 5’PTMCodon) and genomic \(\beta^s\)-globin expression cassettes were co-transfected into 293T cells and total RNA was harvested for analysis by RT-PCR. Primers specific for \textit{trans}-spliced products were used to amplify the repaired RNA, which yielded a DNA product of 410bp. The products were purified from an agarose gel and sequenced to confirm their identity. The transfection performed is listed above each lane in the gel. The term “mix” denotes mix controls that were used to ensure that \textit{trans}-splicing events did not occur during RNA workup and analysis but rather occurred intra-cellularly. Mix controls were made by combining pre-lysis cells from separate PTM and \(\beta^s\) target transfections and then processing these mixed cells similarly to the experimental samples from this point onward. B. The mix samples contain both PTM and \(\beta\)-globin target RNAs. Total RNA from mix samples was reverse transcribed and amplified using either \(\beta\)-globin or PTM specific primers and the products run on agarose gels to confirm that both the 5’PTM and \(\beta^s\)-globin target RNAs were present in the mix sample even though no \textit{trans}-splicing products resulted.
We then examined the efficiency of trans-splicing repair of the βs-globin RNA using a Bsu36I restriction enzyme digest assay. RT-PCR amplification using primers to β-globin’s 5’UTR and exon 2 was followed by an overnight Bsu36I digest. The RT-PCR design allowed for both the cis (βs-globin encoding) and trans-spliced (βwt-globin encoding) products to be amplified equally and the Bsu36I enzyme was used to distinguish between the two (Figure 3.4 A). We then attempted to quantify the levels of repair by densitometric analysis of the resulting restriction fragments. However, we could not detect any digestion fragments using any of the 5’PTMs (Figure 3.4 B) and estimated that the level of repair would need to be 5% or more in order to be identified using this assay (Figure 3.4 C).
Figure 3.4 Quantifying the levels of sickle β-globin repair.

A. Bsu36I restriction enzyme digest used to quantify repair of sickle β-globin RNA. After RT-PCR using primers that bound both repaired and mutant targets, a Bsu36I restriction site existed only in the repaired β-globin product thus allowing for distinction between the two amplification products. B. Amplification products from cells co-transfected with 5'PTM and genomic βs-globin expression cassettes and digested with Bsu36I. The sickle transcript is in significant excess over the repaired RNA. Although trans-splicing is occurring (see Figure 3.3 A), the levels are not high enough for detection via the Bsu36I restriction digest assay. C. The detection limit of the Bsu36I assay is a 5% repair level. Different ratios of digested and undigested normal β-globin RT-PCR products were used to identify the lower detection limit of repaired product in the assay.
A

<table>
<thead>
<tr>
<th>B-Globin</th>
<th>RT-PCR Product Size</th>
<th>Bsu36I Fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wildtype</td>
<td>182bp</td>
<td>115bp &amp; 67bp</td>
</tr>
<tr>
<td>Sickle (B^o)</td>
<td>182bp</td>
<td>Will not digest</td>
</tr>
</tbody>
</table>

B

C

Approximately 100ng total loaded each lane
3.2.4 Discussion

This work has shown that the first exon of a therapeutically relevant gene can be repaired via 5’PTMs and SMaRT technology. It is also the first time SMaRT has been applied to sickle cell anemia and the trans-splicing level achieved illustrates an important general consideration when applying this therapeutic approach. The target’s endogenous splicing rate plays an important role in the success of this trans-splicing therapy and hence should be considered when selecting targets for SMaRT. The β-globin gene is normally spliced very efficiently (Audibert et al., 2002) and perhaps this is why, although we can detect the repaired transcript in the sickle cell case, it is not in high abundance. The sickle cell mutation does not impact the β-globin gene’s ability to splice, but instead leads to a structurally compromised β-chain within the hemoglobin tetramer (Ingram, 1956, 1959; Perutz, 1960). Our 5’PTM constructs may have been unable to compete effectively against this efficiently cis-spliced target.

Our future experiments will attempt to improve repair efficiency of the 5’PTM in two ways. First, we will explore design features that may allow our PTM to compete more effectively against β\(^s\)-globin cis-splicing. Perhaps the 62 nucleotide 5’PTM tag is folding in a manner that prevents the binding domain clear access to the β\(^s\)-globin target (figure 3.5). We could shorten the tag length to the minimum required for 5’PTM-specific primer binding or remove it altogether and use the 5’PTMCodon construct to still clearly distinguish between cis and trans-spliced RNAs. Mansfield and colleagues used
an intron to address exon definition issues, and, although our PTM’s cargo is much smaller than that used in their study, including an intron may still increase the PTM’s attractiveness to the spliceosome (Mansfield et al., 2003). The intron may also provide non-splicing related benefits to the trans-spliced product as was described in section 3.1.

Figure 3.5 Possible RNA secondary structure of the 5’PTM’s tag.
The tag sequence used in the 5’PTMs was folded via the Mfold web server (Zuker, 2003). The predicted structure with the lowest free energy value (deltaG) is shown.

The second option to improve repair efficiencies in the sickle cell case is to somehow reduce the $\beta^s$-globin cis-splicing efficiency. Perhaps a 5’PTM could be designed that artificially destabilizes the $\beta^s$-globin pre-mRNA to create a “stalling” effect. The 5’PTM binding domain could be extended to include binding of the 5’cap
region and since the first exon is defined via interactions between the cap and IVS-1 5’ss (Berget, 1995; Lewis et al., 1996), PTM binding of both components on the target pre-mRNA could potentially further hide the endogenous sequence from the splicing machinery. However, this approach would lengthen the dsRNA generated and therefore we would need to monitor for activity of nuclear dsRNA processes like ADAR (discussed in section 4.2.1) that, if applicable, may impede our trans-splicing levels.

We have shown that 5’PTMs can be used to correct the first exon of a therapeutically relevant gene in HEK293T mammalian cell co-transfection experiments. Using a genomic mutant copy of the β-globin gene, we found that 5’PTMs can trans-splice to repair the defect causing sickle cell anemia at the RNA level (Figure 3.3). This is the first time a 5’PTM has been specifically targeted to correct disease-causing mutations in a gene’s first exon. Further research into our 5’PTM designs may not only improve our repair efficiency level, but may also add to our general understanding regarding splicing characteristics unique to first exons.

3.2.5 Materials and Methods

Plasmids

Three 5’PTM constructs were designed and cloned into the RNA polymerase II expression vector pEGFPN1 (Clontech) using SacI and AgeI restriction enzymes (NEB). The EGFPN1 ORF was then removed by digesting with AgeI and NotI restriction enzymes (NEB), gel extracting the desired EGFPN1-free vector (Qiaquick Gel Extraction
Kit from Qiagen), and filling-in the recessed ends (Klenow large fragment from Fisher Biosciences). T4 DNA ligase (NEB) was then used to bluntly ligate the ends. The EGFPN1 ORF was deleted to avoid read-through transcription events that could potentially generate a PTM/EGFP hybrid transcript. The names of the 5’PTM constructs used were 5’PTM1, 5’PTMCodon, and 5’PTMDISE. The 5’ splice sites used for each 5’PTM were as follows: 5’PTM1 – gtgagt (canonical), 5’PTMCodon - gag gtgagt (sequence found to be highly efficient in β-globin IVS-1 by Roca et al. (Roca et al., 2005)), 5’PTMDISE – gtgggt (hybrid of canonical and endogenous β-globin IVS-1 5’ss).

The β-globin genomic sequences were cloned into the RNA polymerase II expression vector pcDNA3.1(+)Hyg (Invitrogen) using NheI and BamHI restriction enzymes (NEB). Two versions of genomic β-globin constructs were made: wildtype and sickle cell (A to T mutation at nucleotide 70).

All expression plasmids were sequenced for accuracy before use in transfection experiments.

Transfections

Human embryonic kidney (HEK) 293T cells were plated onto either 6-well plates or 100mm dishes and allowed to grow overnight to achieve 40-80% confluency. The media (DMEM (Invitrogen), 10% fetal bovine serum (HyClone), 1x penicillin/streptomycin (Invitrogen)) was replaced the next morning immediately before transfection. A 2:1 ratio of PTM to β-globin expression plasmid was used for all co-transfected samples. 4ug (6 well plates) or 10ug (100mm dishes) of total DNA was
suspended in 150ul Opti-MemI (Invitrogen) and allowed to interact with Superfect lipid transfection reagent (Qiagen) for 10 minutes at room temperature. The DNA to Superfect ratio was 1:3 for all transfections. The DNA/Superfect complexes were then dripped onto the cells and placed in a humidified 37°C incubator with 5% CO2 for 48 hours.

**RNA Harvest and RT-PCR**

48 hours after transfection, cells were trypsinized and total RNA was harvested using Qiagen’s RNeasy Plus Miniprep kit (with a DNase step). The mix controls were made after trypsinization, but before cell lysis and contained 1/3 the cells from each relevant well. The total RNAs were quantified using a NanoDrop spectrophotometer (NanoDrop Technologies).

Equal amounts of each total RNA was used in a one-step RT-PCR reaction with the following controls. The -RT control used total RNA from a co-transfection sample in a reaction containing only Platinum taq (Invitrogen). The water control was a “no template” control for the RT-PCR. We used Invitrogen’s Superscript III reverse transcriptase/Platinum taq polymerase and the following conditions for specifically detecting the trans-spliced product: 56°C for 30 minutes, 94°C for 2 minutes, 25 cycles of (94°C 30 seconds/60°C 30 seconds/72°C 30 seconds) then 7 minutes at 72°C with a final hold at 4°C. The primers used to specifically amplify trans-spliced products bound to the 5’PTM-specific tag and β-globin exon 2. The RT-PCR products were then gel purified using high purity agarose-1000 (Invitrogen) and sequenced.

**Bsu36I Assay**
Total RNA from β\(^8\) / 5’PTM co-transfection experiments was isolated as previously described and amplified via one step RT-PCR using primers that would amplify both *cis* and *trans* spliced products (primers bound 5’UTR and exon 2 of β-globin) using the following conditions: 56°C for 30 minutes, 94°C for 2 minutes, 25 cycles of (94°C 30 seconds/57°C 30 seconds/72°C 15 seconds) then 7 minutes at 72°C with a final hold at 4°C. The RT-PCR products were gel purified, digested overnight with Bsu36I (NEB) and then run on a 3% agarose-1000 (Invitrogen) gel to resolve the fragments.
3.3 Using 5'PTMs to repair the sickle cell mutation in primary proerythroblasts from sickle cell patients.

3.3.1 Summary

Our 5’PTM1 construct was successfully introduced into cultured proerythroblasts from sickle cell patient peripheral blood using a number of delivery methods. These included lipid-based transfection, electroporation, and lentiviral transduction approaches. We were able to detect the repaired product via nested RT-PCR and sequencing while concurrently testing the feasibility of different delivery approaches to these primary cells.

3.3.2 Introduction

After detecting repaired β5-globin in a mammalian cell line model, we next pursued a more relevant setting for our potential therapeutic. Primary cells from sickle cell patients (HbSS) were a logical next step in that the 5’PTM could be tested in an intra-cellular, albeit isolated, disease state. However, since mature red blood cells are enucleated (Geiduschek & Singer, 1979), they are devoid of the transcriptional and splicing activities necessary for expression and trans-splicing. We therefore decided to target HbSS patient primary erythrocyte precursors and specifically the transcriptionally active proerythroblast lineage (Dormer et al., 1981) (figure 3.6).
Hematopoietic stem cells, under the influence of specific growth factors, are directed through the erythrocyte lineage of the myeloid line. High levels of β-globin gene expression characterize the proerythroblast stage (Proerythroblast image courtesy of Dr. Murat Arcasoy, DUMC).

Figure 3.6 Erythropoiesis and the proerythroblast lineage.

- CFU-M,L (self-renewing)
- CFU-S (myeloid line starts)
  - BFU-E
  - CFU-E
- **Proerythroblast** (β-globin expression high)
- Erythroblast (3 stages)
  - Lose nucleus to become...
  - Reticulocyte
  - Erythrocyte
Studies by Panzenbock and Arcasoy both examined growth factor requirements for *in vitro* culturing of the erythropoietic lineage from CD34+ peripheral blood samples (Panzenbock et al., 1998; Arcasoy & Jiang, 2005). Panzenbock and colleagues devised culturing conditions that allowed for significant accumulation of pure erythrocyte precursor populations. This population was also able to differentiate into mature erythrocytes under altered growth factor conditions. Arcasoy and colleagues examined the relative importance of each growth factor involved in culturing CD34+ peripheral blood samples to the proerythroblast stage as well as deciphered which signaling pathways were involved in sustaining a proerythroblast population.

In collaboration with Dr. Arcasoy and Dr. Marilyn Telen, we cultured proerythroblasts from peripheral blood of sickle cell patients obtained during their routine visits to the Duke Comprehensive Sickle Cell Center. Our 5’PTM1 construct was then delivered either as an expression plasmid or via lentiviral vector into these primary sickle proerythroblasts and RNA was later analyzed for repair of the βs-globin transcript.

### 3.3.3 Results

We obtained peripheral blood from sickle cell patients at the Duke Comprehensive Sickle Cell Center (IRB # 894055R16ER), isolated peripheral blood mononuclear cells (PBMC), and isolated CD34+ cells from this population using the MACS CD34 Progenitor Cell Isolation Kit (Miltenyi Biotec). This kit used a monoclonal hapten-conjugated CD34 antibody for selection and an anti-hapten antibody bound to
magnetic beads for magnetic separation. The CD34+ cells were then cultured for 9-10 days using the media described in section 3.3.5.

When the cells reached day 10, we performed cytospins for Wright-Giemsa staining and examined the cells to ensure they had the expected proerythroblast morphology. Figure 3.7 shows cytospin/staining results from two different sickle cell patient samples.

Figure 3.7 CD34+ cells in sickle cell patient peripheral blood can be cultured to proerythroblasts in vitro.
Stained cytospins of day 10 proerythroblasts from two different sickle cell patients. The sample from patient 1 is on the left and patient 2 is on the right. The large nucleus relative to the cytoplasmic space is one hallmark of this cell type.
We also incubated the day 10 cells with a monoclonal antibody to CD71 and performed flow cytometry to ensure the samples were primarily proerythroblasts. CD71 is the transferrin cell surface receptor that is expressed at high levels in erythrocyte precursors (Panzenbock et al., 1998). A representative flow result is shown in figure 3.8.

Figure 3.8 The day 10 cultured cells are primarily erythrocyte precursors.
Flow cytometry using a PE-conjugated anti-CD71 monoclonal antibody revealed a greater than 95% pure erythrocyte precursor population after culturing of the sickle cell patient CD34+ cells in vitro.
We used the same 5’PTM1 construct as was described in section 3.2, but retained the original backbone plasmid’s EGFP ORF (pEGFPN1, Clontech) in order to assess transfection/electroporation efficiencies via fluorescent microscopy. We called this construct 5’PTM1EGFPN1 and it was used as the expression plasmid for all non-viral 5’PTM introduction into the sickle cell primary proerythroblasts.

Lipid-based transfection using Superfect (Qiagen) and electroporation using the Amaxa Nucleofector apparatus was performed on day 9-10 sickle cell patient proerythroblasts. We examined the treated cells under a fluorescent microscope to estimate the efficiency of each delivery method 24 hours after treatment. The lipid-based delivery at that time-point resulted in only 5% GFP positive cells under the microscope whereas electroporating produced around 20% GFP positive cells. The GFP expression plasmid provided as a positive control by Amaxa resulted in 40% GFP positive cells. We waited until 48 hours post-delivery to harvest total RNA and then used 5’PTM and β-globin specific primer sets to determine if both transcripts had been expressed. The results are shown in figure 3.9

We then used trans-splicing specific primers and one-step RT-PCR to determine if the sickle β-globin transcript had been repaired. Trans-spliced product was detected in both transfected and electroporated samples after RT-PCR amplification and agarose gel electrophoresis (figure 3.10 A & B). The 210bp trans-spliced bands were gel extracted and confirmed via sequencing (figure 3.10 C).
Figure 3.9 The 5'PTM1 and β-globin transcripts are both expressed in proerythroblasts cultured from sickle cell patients after transfection or electroporation of 5'PTM1EGFPN1.

A. RT-PCR using 5'PTM1 specific primers produced the expected 350bp fragment after using Superfect (lipid-based reagent) or the Amaxa electroporator (lanes 3 and 6 respectively) to deliver the 5'PTM1EGFPN1 plasmid. Lane 2 is a lane unrelated to the experiment, lane 4 is using the lipid-based reagent on cells alone and lane 5 is electroporation using a plasmid provided with the electroporation kit. B. RT-PCR using β-globin specific primers produced the expected 478bp fragment after using the Amaxa electroporator to deliver the 5'PTM1EGFPN1 plasmid (lane 2) or buffer alone (lane 3). More template was used to produce lane 2 and hence the band is more prominent. Superfect transfections produced similar results (data not shown).
Figure 3.10 5’PTM1 repairs the sickle β-globin transcript in patient proerythroblasts after transfection or electroporation of 5’PTM1EGFPN1.

A. RT-PCR using trans-splicing specific primers produced the expected 210bp fragment after Superfect delivery of 5’PTM1EGFPN (lane 3). Lane 2 is using the lipid-based reagent on cells alone, lane 4 is a water template for RT-PCR and lane 5 is total RNA subjected to PCR alone. B. RT-PCR using trans-splicing specific primers produced the expected 210bp fragment after electroporation of 5’PTM1EGFPN1 (lane 3). Lane 2 is electroporation using a plasmid provided with the electroporation kit, lane 4 is electroporation with buffer alone, lane 5 is a water template for RT-PCR and lane 6 is total RNA subjected to PCR alone. C. The 210bp trans-spliced product in B was gel purified and sequenced in the forward (top) and reverse (bottom, note is reverse complement) directions. The two fragment traces for each direction are from the same scan.
Sickle cell mutation is repaired

FORWARD

Sickle cell mutation is repaired (reverse complement)

REVERSE

Exon 1

Exon 2

Exon 1 (5'UTR) (rev. comp.)

5'PTM1-specific tag (rev. comp.)
Although we were able to deliver our 5’PTM1 expression plasmid to the primary cells using transfection/electroporation, we hoped to improve upon our maximum transfer efficiency to the primary cells (20% achieved using electroporation and estimated by fluorescent microscopy). Hematopoietic stem cells have been shown to be amenable to lentiviral gene therapy (Miyoshi et al., 1999; May et al., 2000; Hanawa et al., 2002) and we, in collaboration with Drs. John T. Gray and Derek Persons at St. Jude Children’s Research Hospital, developed a self-inactivating lentivirus that expressed our 5’PTM1. Our colleagues provided a backbone self-inactivating (SIN) HIV-1 lentiviral transfer vector into which we cloned our 5’PTM1 cassette (figure 3.11 B). The St. Jude backbone vector, pcL20cmLARbV5Dg, was first described in a 2004 publication by Hanawa et al (Hanawa et al., 2004) and its inclusion of β-globin locus control region elements (LCR) was found to improve transgene expression levels in mice. The researchers attributed this to the LCR’s ability to ward off position effects that may have otherwise dampened transgene expression. We removed the γ-globin transgene sequences from this backbone vector and introduced our 5’PTM1 minimal cassette (no EGFPN1 sequences used) under the control of the β-globin promoter and an SV40 polyadenylation signal (figure 3.11 B). After sequence confirmation, we sent the vector to St. Jude for lentiviral production. Since researchers at St. Jude found differences in delivery efficiencies between vsv-g and amphotropic enveloped viruses when applied to hematopoietic stem cells, they provided both pseudotypes for our work. They also provided GFP-expressing lentiviral supernatant
(figure 3.11 A) so that we might monitor these potential differences in transduction efficiencies.

**Figure 3.11 Lentiviral constructs.**

A. GFP-producing self-inactivating (SIN) lentiviral construct. Vsv-g and amphotropic enveloped lentiviruses were made using a 4-plasmid production scheme by the viral vector laboratory at St. Jude Children’s Research Hospital (Memphis, TN) and supernatant was then shipped to us. B. Drs. John T. Gray and Derek Persons at St. Jude Children’s Research Hospital kindly provided the backbone plasmid (pcL20cmLARbV5Dg) necessary for us to construct the vector shown here (backbone plasmid not shown). We cloned a minimal 5’PTM1 sequence under the control of the native β-globin promoter into the backbone plasmid along with an SV40 polyadenylation signal. After sequence verification, we sent this construct to St. Jude where vsv-g and amphotropic enveloped SIN lentiviruses were produced. Concentrated supernatant was then sent back to our laboratory.
Figure 3.12 describes the first of two lentiviral transduction approaches we used to deliver the 5’PTM1 into erythrocyte precursors cultured from sickle cell patient peripheral blood. In this scheme, we cultured the CD34+ isolated cells all the way to the proerythroblast stage prior to infecting the sample with our 5’PTM1 lentivirus.

**Lentiviral Transduction Scheme I**

Isolate CD34+ from HbSS patient blood

![Diagram of the transduction process]

- Culture cells for 9-10 days to proerythroblast stage
- Infect proerythroblasts with lentivirus
- Harvest cells after 48 hours for imaging, flow cytometry and RT-PCR

*Figure 3.12 Lentiviral transduction scheme I.*

The procedure used to deliver the lentivirus to sickle cell primary proerythroblasts. Unlike in transduction scheme II (to be discussed later), the cells are infected with the lentivirus after they have reached the proerythroblast stage.

48 hours post-infection, we harvested the cells and used those transduced with the GFP lentivirus in a flow cytometry experiment to estimate our delivery efficiency into the proerythroblasts. Figure 3.13 shows that GFP+ cell counts increased as greater volumes of lentivirus were used per infection. The maximum lentiviral supernatant volume used to infect the cells in these experiments was 100ul and this corresponded to approximately 17% of cells expressing GFP (figure 3.13 A). We also wanted to ensure that our
lentivirus was infecting the desired proerythroblast cell type hence we also stained the cells with a PE-conjugated CD71 antibody. As is shown in figure 3.13 B, the vast majority of cells are proerythroblasts (left) and the GFP+ cells (from 100ul lentiviral supernatant) are proerythroblasts (right). Figure 3.13 C is a plot showing just the GFP data from the cells in B. The top and bottom are PE-CD71 and PE-isotype control stained cells respectively and only the top was quantified and found to be around 19% GFP+ cells.

We then wanted to see if our 5’PTM1 lentivirus was expressing the PTM in the proerythroblasts. We harvested total RNA from cells 48 hours after 5’PTM1 lentiviral infection and performed RT-PCR amplification using 5’PTM1 specific primers (figure 3.14). We detected the transcript after using 100ul 5’PTM1 vsv-g or amphotropic virus, and this RT-PCR product increased when using 500ul of the supernatant (lanes 3 and 9). However, a media exchange performed 24 hours post-infection did not seem to increase transcript production and in fact appeared to diminish it (lanes 4 and 5).
Figure 3.13 The vsv-g GFP lentivirus will infect primary proerythroblasts from sickle cell patients. A. Infecting primary cells with increasing amounts of vsv-g enveloped GFP lentivirus resulted in increased numbers of GFP positive cells upon flow cytometry. The same gate was used in all experiments and $10^4$ events were counted for each group. The M1 and M2 bars were also identical for all experiments. The y-axis label is “counts” and the range is 0 to 200 (bottom to top) for all graphs. B. The GFP positive cells are proerythroblasts. This is shown by PE-CD71 staining on the left (an erythrocyte precursor marker) and concurrent detection of both GFP and PE on the right. The isotype control experiments are shown below. The same gate was used in all experiments and $10^4$ events were counted for each group. The M1 and M2 bars were also identical for all experiments. 100ul virus was used. C. Histogram of GFP data in B presented separately. The top is from the PE-CD71 stained cells and the bottom is from the PE-Iso stained cells. Only the top histogram was quantified and the GFP level was around 19%. The y-axis label is “counts” and the range is 0 to 200 (bottom to top) for both graphs.
A

No virus
<1% GFP positive cells

10ul virus
2% GFP positive cells

50 ul virus
12% GFP positive cells

100ul virus
17% GFP positive cells
Figure 3.14 The 5’PTM1 transcript is expressed from vsv-g or amphotropic enveloped lentiviruses within sickle cell patient proerythroblasts.

RT-PCR using 5’PTM1 specific primers produced the expected 350bp product following administration of 100ul or 500ul of either 5’PTM1 lentivirus (lanes 2, 3, 8, 9). The faint band around 290bp is the result of non-specific priming. Mock lanes are untreated proerythroblasts and lanes 4 and 5 are infected cells that underwent a media exchange after 24 hours. Lanes 10 and 11 are controls for the RT-PCR reaction.

We used RT-PCR with trans-splicing specific primers to see if we could detect the repaired sickle β-globin transcript after infection with the vsv-g 5’PTM1 lentivirus and the results are shown in figure 3.15. Both 500ul and 1000ul supernatant applications led to the expected 219bp trans-spliced RT-PCR product (figure 3.15 A) and we sequenced the band in lane 3 in both forward and reverse directions (figure 3.15 B). The sequence traces show the repaired nucleotide in each strand as well as the exon 1/exon 2 β-globin junction and 5’PTM1-specific tag sequence. This data indicated that a 5’PTM1-driven repair of sickle β-globin pre-mRNA had occurred in proerythroblasts cultured from sickle cell patient peripheral blood.
Figure 3.15 Lentiviral delivery of 5′PTM1 into sickle cell patient proerythroblasts results in repair of the sickle cell transcript.

A. Nested RT-PCR using trans-splicing specific primers produced the expected 219bp product following administration of 500ul or 1000ul of vsv-g 5′PTM1 lentivirus to the primary cells (lanes 3 and 4). Lane 2 is uninfected cells, lane 5 is a water template in the original RT-PCR, lane 6 is a reverse transcriptase-free experimental sample in the original RT-PCR and lane 7 is a water template used in the nested PCR reaction. B. Sequence scans of the 219bp fragment seen in part A, lane 3. Forward and reverse sequence scans are shown. Forward: Sequencing in the forward direction. Left is corrected nt 70. Right is exon 1/exon 2 junction from same sequence scan. Reverse: Sequencing in the reverse direction hence all is reverse complement. Left is corrected nt 70. Right is exon 1/5′PTM1-specific tag junction from same sequence scan.
We attempted to use an RNase protection assay to directly detect the repaired RNA, but levels of trans-spliced product or even 5’PTM1 expression were not high enough to see using this assay (figure 3.16). However, our RPA detection efforts were somewhat limited by the amount of input RNA we could test. Prior 5’PTM/β² 293T co-transfection experiments produced false positive results after using large quantities of sample RNA in the RPA (figure 3.17). Specifically, a “positive” band appeared in the mix control RNA. The mix control was total RNA from HEK293T co-transfection experiments where 5PTM1 and sickle β-globin expression plasmids were separately transfected into 293T cells then mixed together during cell harvest (prior to lysis). The 5’PTM1 and sickle β-globin RNAs were therefore produced in separate cells, but now joined in one tube. Incubation of the RPA probe with this mixed total RNA should not produce the protected 224bp double-stranded RNA band, but it began to appear as the input RNA levels increased (figure 3.17). Although this false positive was uncovered in a 293T experiment, we felt a similar scenario was possible in our primary cells since both the RPA probe and protocol used were the same. We hypothesized that perhaps the 5’PTM1 and sickle β-globin RNAs were folding in such a way so that one RPA probe could bind both simultaneously. This is plausible since the 5’PTM1 and sickle β-globin transcripts share a nearly identical exon 1 sequence (save for the point mutation at nucleotide 70).
Figure 3.16 RNase Protection Assay.
A. Probe binding sites to individual RNAs with the RNA probe in red. After probe binding and RNase treatment, the dsRNA protected fragment sizes are as described to the right each graphic. A symbol (in parenthesis) is used to identify the respective fragment in the gel in part B. The loop seen in the probe binding to $\beta^s$ indicates the point of non-binding at the sickle cell nucleotide. B. A representative RPA result. No trans-spliced fragment is seen in the experimental lanes (lanes 7 and 8). Lane 1 is the undigested purified probe alone. Bands below the 254bp band in lanes 1, 2 and 3 are from probe degradation. Control 1 is probe and buffer alone, control 2 is probe and non-specific yeast RNA, and control 3 is the result of treating control 2 with RNase. Mock indicates patient proerythroblasts without lentiviral infection. Lanes 7 and 8 are the 5'PTM1 amphotropic lentiviral infections of patient proerythroblasts. No trans-splicing is detected in either of these lanes. Mix indicates a negative control used in the HEK293T 5PTM1EGFPN1/$\beta^s$ co-transfection experiments and is used here to show the location of the 5'PTM1 protected fragment.
Figure 3.17 Possible false positive result in the RPA experiment when using high levels of input RNA. All data here is from HEK293T co-transfections using expression plasmids (see section 3.2), but illustrates the potential false positive result when a single RPA probe binds both 5’PTM1 and sickle β-globin transcripts simultaneously (potentially due to a combination of homology and high levels of both transcripts in the RPA) (lane 7). All labels are as described in figure 3.16, but with lanes 6-10 from plasmid co-transfection into 293T cells. The RPA probe and protocol used here are identical to that used in figure 3.16.
After meeting with our collaborators at St. Jude, we altered our viral transduction protocol in an attempt to improve our transduction efficiencies. Lentiviruses, unlike their oncoretroviral counterparts, are able to infect non-dividing cells and do so via preintegration complex (PIC) interactions with the nuclear pores (Bukrinsky et al., 1992; Naldini et al., 1996). Nuclear pores in proerythroblasts and other erythrocyte precursors become clustered as the nucleus condenses in preparation for enucleation (Krauss et al., 2005) and perhaps this was impeding our 5’PTM1 lentivirus from efficiently entering the nucleus under scheme I (figure 3.12). We therefore modified our transduction protocol so that infection occurred at the CD34+ stage, very early in erythropoiesis, when nuclear pores are more evenly distributed around the nucleus (Krauss et al., 2005). This modified approach is outlined in figure 3.18 and was termed “lentiviral transduction scheme II”.

Lentiviral Transduction Scheme II

Isolate CD34+ from HbSS patient blood

Prestimulate CD34+ cells for 48 hours and prepare Retronectin plates with lentivirus (4-6 hours prior to addition of cells).

After prestimulation, spin down cells and resuspend in fresh prestimulation media. Add cells to Retronectin-coated plates preincubated with lentivirus. Incubate for 12 hours.

After transduction, spin down cells and then culture in proerythroblast media for 9 days.

Harvest cells for flow cytometry, RT-PCR or RPA analysis

Figure 3.18 Lentiviral transduction scheme II.

The procedure used to deliver the lentivirus to sickle cell primary proerythroblasts as suggested by our collaborators at St. Jude Children’s Research Hospital. This protocol differs from transduction scheme I in that infection occurred at the CD34+ cell stage and included a pre-stimulation step prior to introduction of the virus (see Methods for details). Also, Retronectin-coated plates were used in place of polybrene to further improve transduction efficiency. HbSS denotes patients homozygous for the sickle cell mutation.
We tested this modified approach by infecting pre-stimulated CD34+ cells from sickle cell patients using either the vsv-g or amphotropic enveloped GFP-expressing lentiviruses. After allowing for differentiation to the proerythroblast stage, we harvested and stained the cells with PE-CD71 then counted the number that were both GFP and PE positive via flow cytometry (figure 3.19). We also captured images using fluorescent microscopy and these are shown below the flow data. We found a significant increase in GFP positive cells versus what was detected using transduction scheme I (figure 3.13) with either pseudotyped virus, however we used more supernatant in scheme II so it is difficult to make direct comparisons (100ul maximum versus 700ul minimum vsv-g GFP lentivirus in schemes I and II respectively).

One interesting finding from scheme II was the higher transduction efficiency when using the vsv-g envelope (67% GFP positive cells) versus the amphotropic lentivirus (47% GFP positive cells) when all gates, M1/M2 bars and total number cells counted were kept constant between the two (figure 3.19). It is also notable that the proerythroblast cell type was quite pure in these populations (figure 3.19, greater than 99% CD71+ cells in both), suggesting that cytokine stimulation and early viral infection did not detrimentally affect differentiation to the proerythroblast stage.
Figure 3.19. Viral transduction scheme II leads to high lentiviral GFP expression in proerythroblasts and the viral envelope may have an impact on transduction efficiency.

All data in A and B are from the same sickle cell patient. A. 1000 μl of vsv-g GFP lentivirus was used to infect pre-stimulated CD34+ cells and then cultured to the proerythroblast stage. Cells were then harvested for flow cytometry using a PE-CD71 antibody and assayed for both PE and GFP expression. The forward/side scatter plot shows the cells that were gated and used for all analyses. The same M1 and M2 bars were used for all analyses. The PE-isotype controls were less than 1% PE+ (data not shown). Cells were also photographed using a fluorescent microscope (shown below the flow cytometry data). B. Same as part A, but using 1000 μl of amphotropic GFP lentivirus. Gates and M1/M2 bars are the same as those used in part A.
5’PTM1 vsv-g and amphotropic enveloped lentiviruses were then transduced using scheme II and total RNA was tested for 5’PTM1 expression via RT-PCR (figure 3.20). The PTM transcript was expressed within the HbSS primary cells using either pseudotyped lentivirus (lanes 5 and 6) and RT-PCR products were sequenced to confirm their identities. The total RNA from 5’PTM1 lentiviral infections using scheme II was then assayed for trans-splicing using nested one-step RT-PCR followed by sequencing. Figure 3.21A shows the repaired 400bp product after delivery of either pseudotyped 5’PTM1 virus. The sequence scans of one repair product is shown in part B. Figure 3.21C is the result of our attempt to eliminate a possible particular false positive result and we did not detect the repaired product in this experiment. This may have been due to the stringent purification conditions used in the assay and this will be explored further in the Discussion section.
Figure 3.20 The 5’PTM1 transcript is expressed from vsv-g or amphotropic enveloped lentiviruses within sickle cell patient proerythroblasts using viral transduction scheme II.

RT-PCR using 5’PTM specific primers produced the expected 337bp product following administration of 1000μl 5’PTM1 vsv-g or amphotropic enveloped lentivirus (lanes 5 and 6). Lane 2 is untreated proerythroblasts, lanes 3 and 4 are GFP-expressing lentiviruses and lanes 7 and 8 are controls for the RT-PCR reaction.
Figure 3.21 5’PTM1 lentiviral delivery into sickle cell patient proerythroblasts using transduction scheme II results in repair of the sickle cell transcript.

A. Nested RT-PCR using trans-splicing specific primers produced the expected 400bp product following administration of 1000ul of vsv-g or ampho 5’PTM1 lentivirus to the primary cells (lanes 5 & 6). The marker band is 400bp. Lane 2 is uninfected cells, lanes 3 and 4 are GFP viruses, lane 7 is a reverse transcriptase-free experimental sample in the primary RT-PCR and lane 8 is a water template used in the nested PCR reaction. B. Sequence scans of the 400bp fragment seen in part A, lane 6. Analysis is as in figure 3.15B. C. Attempting to eliminate the possibility of a particular false positive result. Two-step RT-PCR using a streptavidin/biotin purification step (see Methods for details) did not result in detectable 555bp trans-spliced product after agarose gel electrophoresis (lanes 5 and 6). The positive control was total RNA from a plasmid-based trans-splicing event in HEK 293T cells and the expected 555bp band is seen (lane 7). The templates for lanes 2-6, 8 are as described in part A with the exception of H20 & H20(2), which are water templates in standard PCR tubes used in the RT and PCR reactions respectively. The three bands in the marker lanes are from top to bottom: 700bp, 600bp (darkest band), and 500bp.
3.3.4 Discussion

We have shown that proerythroblasts can be cultured \textit{in vitro} from sickle cell patient peripheral blood (figures 3.7 and 3.8) and that our 5’PTM1 construct can be delivered to these cells by multiple means (figures 3.9, 3.14, and 3.20). We also detected the repaired \textit{trans}-spliced product in all cases via RT-PCR and sequencing (figures 3.10, 3.15, 3.21).

Since a “mix” control like that used in the 293T co-transfection experiments (section 3.1) was not an option for primary cell assays because of their endogenous target expression, we relied on thorough sequencing to verify our results. Although this was sufficient to validate the repair products, we still pondered how the one-step RT-PCR reactions might potentially generate false positive data. Previous 293T RPA experiments had uncovered a false positive potential stemming from the use of large amounts of RNA and close similarities between mutant and wildtype β-globin exon 1 sequences (figure 3.17). We initially used this discovery to avoid testing large amounts of proerythroblast RNA in our primary cell RPA experiments since we anticipated that the \textit{trans}-spliced product levels would likely be dwarfed by the levels of each substrate produced (PTM and β-globin transcripts). If the probe were able to bind both substrates simultaneously, it would be difficult to distinguish true from false positive RPA results. Thus, although our initial attempts to detect the \textit{trans}-spliced product in primary cells via RPA were
unsuccessful (figure 3.16), we made minimal alterations and maintained fairly low RNA input levels.

The RPA experience also shed light on how similarity in exon 1 could theoretically lead to a false positive result in a one-step RT-PCR reaction. For example, if the sickle cDNA were to act as a primer to the repaired cDNA’s second strand in the PCR portion of a one-step RT-PCR reaction, a product could arise that contained both the 5’PTM-specific tag and the sickle nucleotide. We felt our mix control (plus sequencing) in the 293T experiments was sufficient to monitor for this (sections 3.2 and 3.4) and the forward/reverse sequence data verified the one-step RT-PCR primary cell data. However, we wanted to try to eliminate even the potential for this theoretical false positive from our analysis. We therefore devised a two-step RT-PCR protocol incorporating a streptavidin/biotin-based purification step (described in detail in the Methods section) to remove all cDNAs lacking a 5’PTM-specific tag prior to beginning the PCR amplification. This would remove the sickle cDNA from the PCR reaction and therefore prevent the false positive. However, this multi-step procedure may have also resulted in a considerable loss of RNA, which appeared to be the case from the barely detectable positive trans-splicing control band seen in figure 3.21C. Furthermore, the streptavidin/biotin purification step may have been stringent enough to remove a substantial amount of truly repaired cDNAs. These factors may have contributed to our inability to detect the trans-spliced product after using the two-step RT-PCR protocol.
As mentioned previously, we carefully analyzed the relevant sequence data from the primary cell one-step RT-PCR reactions to ensure that the 5’PTM1-specific tag and β-globin exon 1/exon 2 splice junction were both present and that the repaired nucleotide could be clearly detected from both forward and reverse sequencing directions (figures 3.10, 3.15, and 3.21). This combined sequence data confirmed that the products truly stemmed from repaired β-globin transcripts.

The primary cell studies also produced interesting findings regarding lentiviral delivery to proerythroblasts. It is difficult to directly compare the efficiencies between transduction schemes I and II since 100ul and 700ul were the highest and lowest supernatant volumes used in I and II respectively. However, it is clear that scheme II provided robust transduction results (67% GFP at max supernatant volume). Furthermore, improved delivery using scheme II over scheme I supports the theory that lentiviruses may be impeded by clustered nuclear pores like those seen in late pre-erythrocytes.

The differing scheme II transduction efficiencies generated using vsv-g versus amphotropic enveloped GFP lentiviruses was intriguing because this result appears to contradict what Hanawa and colleagues previously reported (Hanawa et al., 2002). They found that infecting pre-stimulated CD34+ cells from peripheral blood using amphotropic rather than vsv-g enveloped lentiviruses led to higher GFP transgene levels using flow cytometry. Our studies were dissimilar in some significant ways though. The Hanawa findings came from non-sickle cell healthy volunteers that had been dosed with G-CSF
prior to blood draw to increase hematopoietic stem cell production. Perhaps sickle CD34+ cells have unique characteristics that make them more susceptible to vsv-g mediated entry and this produced a difference in our study results. The cells in the Hanawa study were also collected for flow cytometry 48 hours post-infection whereas we cultured our cells for 9 days prior to flow analysis. Perhaps the vsv-g enveloped virus was better able to infect cells under our proerythroblast growth conditions and this also contributed to the differing conclusions.

3.3.5 Materials and Methods

Plasmids

5’PTM1 was used as described in 3.2.5 with one change. In order to monitor transfection efficiency of the plasmid into the hemoglobin SS patient proerythroblasts when either lipid transfection or electroporation was used, we preserved the EGFPN1 ORF in the original backbone plasmid (pEGFPN1 from Clontech). We called this version of the 5’PTM1 plasmid “5’PTM1EGFPN1” (note: the lentiviral vector described later did not contain the EGFPN1 ORF).

Isolation of CD34+ Cells from Hemoglobin SS Patient Peripheral Blood.

25-50mL of peripheral blood from hemoglobin SS (HbSS) patients was obtained from the Duke Sickle Cell Clinic (IRB # 894055R16ER) and collected in anticoagulant citrate dextrose solution A (ACD-A) tubes. The blood was diluted 3-4 fold using 0.6% ACD-A in PBS to a final volume of 50mL. 15 mL Lymphoprep (Nycomed Pharma) was
slowly introduced under the blood layer and the tubes were then spun in a swinging bucket rotor at 400xg for 35 minutes at 25°C. The peripheral blood mononuclear cells were then transferred to new tubes and washed three times using 0.6% ACD-A in PBS each time and spinning for 10-15 minutes between washes. The final cell pellet was resuspended in 300ul of cold 0.6% ACD-A plus 0.5% BSA in PBS. The cells were then separated using a pre-separation filter (Miltenyi Biotec) prior to the next step.

The cells were then labeled with 100ul of monoclonal hapten-conjugated CD34 antibody for 15 minutes at 6-12°C using the MACS CD34 Progenitor Cell Isolation Kit (Miltenyi Biotec). 100ul of a blocking reagent was also added. The cells were then washed using cold 0.6% ACD-A plus 0.5% BSA in PBS, spun down, and resuspended in 400ul of fresh wash buffer. 100ul of an anti-hapten microbead slurry was then added and allowed to bind at 6-12°C for 15 minutes. The cells were washed as before and resuspended in 500ul of fresh wash buffer. The labeled cells were separated using a pre-separation filter (Miltenyi Biotec) and added to a column for positive selection of CD34+ cells using magnetic separation.

**Culturing of CD34+ Cells to the Proerythroblast Stage.**

The CD34+ patient cells were placed in a vented T-25 flask with 5 mL proerythroblast media and cultured for 9-10 days in a humidified 37°C incubator with 5% CO2. A partial media exchanged was performed every day for the first 5 days and then every 2 days after that. The proerythroblast media contained 15% fetal calf serum, 0.016% sodium bicarbonate, 0.0007% 2-beta-mercaptoethanol, 0.125 mg/mL.
holotransferrin, 100 units/mL penicillin/streptomycin, 1% BSA, 1 unit/mL erythropoietin, 100 ng/mL human stem cell factor, 40 ng/mL human insulin growth-like factor 1, 10^{-6} mM dexamethasone, and 10^{-6} mM β-estradiol in Dulbecco’s Modified Eagle Media. Proerythroblast identity was confirmed by morphology after cytospin staining and by CD71 labeling in flow cytometry.

**Cytospin Staining and Flow Cytometry**

1x10^6 cells were collected onto glass slides via room temperature centrifugation. The cells were then stained using Wright-Giemsa stain and examined under the microscope to confirm their identity.

For flow cytometry, at least 3x10^5 cells were pelleted and resuspended in phosphate buffered saline (PBS) containing 4% fetal bovine serum (FBS). The cells were then incubated with R-Phycoerythrin (PE)-conjugated monoclonal anti-human CD71 mouse IgG1 antibody (Immunotech) or the PE-conjugated mouse IgG1 isotype control antibody for 1 hour in the dark at 4°C. The stained cells were washed twice in PBS with 4% FBS and then fixed in PBS containing 1% formaldehyde and stored in the dark 4°C until use. Flow cytometry was performed on a FACSCalibur machine (Becton Dickinson) and counts were taken for both GFP (if GFP lentivirus was used) and PE positive cells. All flow data was analyzed using CellQuest software (Becton Dickinson).

**Transfection (Lipid-Mediated) and Electroporation of Primary Cells**

We transfected our 5’PTM1EGFPN1 expression plasmid into HbSS patient proerythroblasts using two different strategies. We first used a lipid-based Superfect
transfection protocol (Qiagen). Various amounts of 5’PTM1EGFPN1 (ranging from 4ug for 6 well plates to 20ug for 100mm dishes) were resuspended in 200ul Opti-MemI (Invitrogen) and incubated with Superfect for 10 minutes at room temperature. The DNA/lipid complexes were then added to either 2-3x10^6 (for 6 well plates) or 3.5x10^7 (for 100mm dishes) HbSS primary proerythroblasts in fresh proerythroblast media. The transfected cells were then placed in a 37°C humidified incubator with 5% CO₂ for 48 hours.

A second strategy was to use electroporation for delivery 5’PTM1EGFPN1 to the primary cells. We used Amaxa’s Nucleofector apparatus (kindly provided by H. Kim Lyerly) and the company’s CD34+ buffer kit for the electroporation. This kit also provided the electroporator settings for use during the electroporation. We electroporated either 4ug of the positive control plasmid provided in the kit (pMAXGFP) or 4ug of our 5’PTM1EGFPN1 plasmid into 5x10^6 HbSS primary proerythroblasts, allowed the electroporated cells to recover in calcium-free DMEM for 10 minutes, then added 1.5mL fresh proerythroblast media to each tube and transferred all to 12 well plates. The transfected cells were then placed in a 37°C humidified incubator with 5% CO₂ for 48 hours.

5’PTM1 Lentiviral Production

A 5’PTM1Lenti fragment was cloned into a backbone vector generously provided by Dr. John T. Gray at St. Jude Children’s Research Hospital (Memphis, TN). The original backbone plasmid, pCL20cmLARbV5Dg, contained a modified γ-globin ORF
under the control of the β-globin promoter and chromosome 11 locus control region (LCR). We removed the promoter and γ-globin sequences by digesting the vector with SbfI and EcoRI restriction enzymes (NEB). We created a 5’PTM1 insert via overlapping PCR that incorporated the β-globin promoter upstream and a SV40 polyA signal downstream of the 5’PTM1 minimal fragment. This insert was cloned into the TOPOTA vector (Invitrogen) and transformed into Top10 Chemically Competent *E. coli* (Invitrogen). Resultant TOPOTA colonies were screened for the 5’PTM1 insert via PCR and positive clones were grown overnight as liquid cultures. DNA was isolated and digested with SbfI and EcoRI then run on an agarose gel. The SbfI/EcoRI 5’PTM1 insert was gel extracted (QiaQuick Gel Extraction Kit, Qiagen) and cloned into the previously digested backbone vector using T4 DNA Ligase (NEB). The ligation was transformed into Top10 Chemically Competent *E. coli* (Invitrogen), plated, screened, and grown overnight in liquid culture. The plasmid DNA was then isolated, sent for sequence confirmation, and re-named pcL20cmLAR5PTM1 before being sent to St. Jude for viral production. Both vsv-g and amphotropic enveloped lentiviruses were made at St. Jude and frozen supernatant was then sent back to the laboratory.

**Viral Transduction I: Transduction of Hemoglobin SS Patient Cells at the Proerythroblast Stage.**

CD34+ cells were isolated from HbSS patient peripheral blood as described previously. These cells were then immediately introduced to proerythroblast media and allowed to differentiate for 9-10 days in a humidified 37°C incubator with 5% CO₂ to
reach the proerythroblast stage. During the incubation period, a partial media exchanged was performed every day for the first 5 days and every 2 days after that.

After cells reached the proerythroblast stage, lentiviral transduction was performed in fresh proerythroblast media using 6ug polybrene (Sigma) per 1mL volume transduction. Different amounts of lentivirus were used for different experiments. After 24 hours, fresh proerythroblast media was added to each well (media addition, not exchange). 48 hours post-transduction cells were harvested and stained with PE-CD71. Flow cytometry was used to verify the proerythroblast cell type (via PE-CD71) and to determine transduction efficiency (via GFP when GFP-expressing lentivirus was used).

**Viral Transduction II: Transduction of Hemoglobin SS Patient Cells at the CD34+ Stage.**

This transduction protocol was similar to that described by Hanawa and colleagues in 2002 (Hanawa et al., 2002). After CD34+ cells were isolated, they were prestimulated for 48 hours before viral transduction. The prestimulation conditions included serum-free media, 300 ng/mL human stem cell factor, 300 ng/mL human Flt3 ligand, 50 ng/mL recombinant human interleukin-6, and 1% bovine serum albumin. The cells were prestimulated for 48 hours in a humidified 37°C incubator with 5% CO₂ at a concentration of 1-2 x 10⁵ cells per mL of the media.

After 48 hours, the cells were collected via centrifugation and resuspended in 2 x 10⁵ cells per mL fresh prestimulation media and added to 20 ug/cm² Retronectin-coated
plates (Takara) that had been previously loaded for 4-6 hours at 37\(^{0}\)C with lentiviral supernatant. The amount of lentivirus used varied between experiments.

The transduced cells were incubated for 12 hours in a humidified 37\(^{0}\)C incubator with 5% CO\(_2\) and then collected by centrifugation. The cells were then resuspended in proerythroblast media and allowed to differentiate for 9-10 days in a humidified 37\(^{0}\)C incubator with 5% CO\(_2\) to reach the proerythroblast stage. A partial media exchanged was performed every day for the first 5 days and then every 2 days after that. Flow cytometry was performed using GFP-expressing versions of the lentivirus along with CD71 staining to ensure that transduction was occurring and doing so in the desired cell type. Proerythroblast identity was further confirmed by morphology after cytospin staining.

**RNA Harvest**

48 hours post transfection/electroporation/viral transduction scheme I, total RNA was isolated either using TrizolLS reagent (Invitrogen) followed by DNaseI treatment (Promega or Roche) or via Qiagen’s RNeasy Plus Total RNA Extraction kit that contains a spin-column based DNA removal step. All RNA was resuspended in RNase-free water. For viral transduction scheme II, total RNA was isolated when the infected cells reached the proerythroblast stage.

**Reverse Transcription and PCR**

Total RNA was either reverse transcribed separately using Superscript III reverse transcriptase (Invitrogen) or in a one-step RT-PCR reaction using a Superscript III RT/
Platinum taq kit (Invitrogen). If cDNA was made in a separate step, the reaction was cleaned using Qiagen’s QiaQuick PCR purification kit prior to use in the PCR step. Platinum taq the enzyme used in all PCR reactions.

For the non-viral one-step RT-PCR reactions, 0.15ug to 1ug of total RNA was used with the following controls. The -RT control contained total RNA from a 5’PTMEGFPN1 sample along with all other RT-PCR components except for the reverse transcriptase. The water control was a “no template” control for the RT-PCR. We used the following conditions for specifically detecting the trans-spliced product: 55°C for 30 minutes, 94°C for 2 minutes, 35 cycles of (94°C 45 seconds/68°C 45 seconds) then 7 minutes at 72°C with a final hold at 4°C. The primers used to specifically amplify trans-spliced products bound to the 5’PTM-specific tag and β-globin exon 2 or 3. We also used primers sets specific to either 5’PTM1EGFPN1 or β-globin under the following conditions to verify the presence of each transcript in the total RNA samples: 55°C for 30 minutes, 95°C for 2 minutes, 25 cycles of (95°C 30 seconds/55°C 30seconds/72°C 30 seconds) then 7 minutes at 72°C with a final hold at 4°C.

For the viral one-step RT-PCR reactions, 0.5 - 1.0ug of total RNA was used with the same controls as described above for the non-viral case. We used the following conditions for specifically detecting the trans-spliced product: 56°C for 30 minutes, 94°C for 2 minutes, 25 cycles of (94°C 30 seconds/56°C 30 seconds/72°C 30 seconds) then 7 minutes at 72°C with a final hold at 4°C. The primers used to specifically amplify trans-spliced products bound to the 5’PTM-specific tag and β-globin exon 3. We then used
1/10 of this product and performed nested PCR using the same forward primer and a reverse primer that bound β-globin exon 2 for 20 more cycles. We also used primers sets specific to either 5’PTM1EGFPN1 or β-globin under the following conditions to verify the presence of each transcript in the total RNA samples: 56°C for 30 minutes, 94°C for 2 minutes, 25 cycles of (94°C 30 seconds/56°C 30 seconds/72°C 30 seconds) then 7 minutes at 72°C with a final hold at 4°C.

For the viral two-step RT-PCR reactions (used only with viral transduction scheme II), 0.4-1.8ug total RNA was reverse transcribed (Superscript III, Invitrogen) using a primer that bound β-globin exon 3 and purified (QiaQuick PCR Purification kit, Qiagen). 80ng of each purified cDNA was then annealed to 5-10 pmol of a biotinylated 5’PTM-specific forward primer by adding salts/high heat then allowing the mixture to cool down slowly on the benchtop to 25°C. When cooled, water was added to bring the final volume to 50ul and the mixture was transferred to streptavidin-coated PCR tubes (Roche). The biotin/streptavidin interaction was allowed to form at 37°C for 3 minutes before all liquid was carefully removed from the coated tube. The tubes were then gently washed 5 times with a buffer containing 0.01M Tris-HCl pH7.5, 0.2M LiCl, and 1mM EDTA (pH 8.0) and all liquid was carefully removed after the final wash. The PCR step was then performed in these same tubes using trans-splicing specific primers under the following conditions: 95°C for 5 minutes, 30 cycles of (95°C 30 seconds/59°C 30 seconds/72°C 45 seconds) then 7 minutes at 72°C with a final hold at 4°C. The -RT and water controls were made as previously described although both were now in
streptavidin-coated PCR tubes. A second water control in a standard PCR tube was also used.

All RT-PCR products were analyzed via agarose gel electrophoresis, gel purified, and confirmed via sequencing.

**RNase Protection Assay**

Ambion Company’s RPAIII kit was used for all assays.

Probe and RNA Co-Precipitation/Hybridization: Identical amounts of total RNA (in ug) were diluted in 80ul of water per tube. The quantity of total RNA used in each experiment varied, but an average was 2.5ug. Three controls were used during the study: no RNA, yeast RNA without RNase digestion, and yeast RNA with RNase enzymes added in levels equal to that used in the experimental samples. The purpose of each control was to assess the integrity of the full-length probe, verify the probe’s specificity for its target, and determine the extent of RNase digestion respectively. Each RNA sample was co-precipitated with either 2, 5 or 10 fmol of radiolabeled gel purified full length RNA probe using 0.5M ammonium acetate and 100% ethanol. The tubes were placed in an -80°C freezer for 20 minutes then the RNA was pelleted using a 4°C microcentrifuge set at 13,000 RPM for 20 minutes. The supernatant was removed and the pellet was allowed to air-dry for 5 minutes. The co-precipitated probe and RNA sample was resuspended in 10-20ul of the hybridization buffer provided in the RPA III kit and allowed to hybridize overnight at either 42°C or 56°C.
RNase Digestion of Unbound RNA: The RNA that did not hybridize to the probe was digested using a combination of RNase A and RNase T1. Various ratios of enzyme to digestion buffer (provided in the RPAIII kit) were used and these ranged from 1:25 to 1:150. The digestion was allowed to proceed for 30 minutes at 37°C and stopped using the RPAIII kit’s RNase Inactivation and Precipitation solution. The samples were ethanol precipitated, resuspended in 6ul Gel Loading Buffer II (provided in the RPAIII kit), and loaded onto a 6% denaturing polyacrylamide gel. The gel was then dried using vacuum and heat (80°C) for 2 hours and exposed overnight on a phophor screen. The screen image was then captured using a Storm Phosphorimager (Molecular Dynamics/GE Healthcare) and the results were analyzed using ImageQuant software (Molecular Dynamics/GE Healthcare).
3.4 Using 5’PTMs to repair β-thalassemia mutations

3.4.1 Summary

We hypothesized that target transcripts with diminished cis-splicing capabilities might be well-suited for SMaRT repair since the technology relies on trans-splicing prevailing over endogenous cis-splicing. We tested this hypothesis in a mammalian cell model. 5’PTM expression plasmids were co-transfected into HEK293T cells along with genomic β-globin expression plasmids carrying mutations that disrupted normal cis-splicing. Each of the β-globin splicing mutants had been shown previously to cause β-thalassemia in certain patient populations. Our results suggest that splicing-impaired pre-mRNAs may indeed be ideal targets for SMaRT.

3.4.2 Introduction

Since SMaRT depends on the PTM’s ability to compete with cis-splicing for spliceosomal components, we postulated that perhaps the inability of the 5’PTMs to repair more of the β⁺-globin RNA was due to the highly efficient cis-splicing of the sickle β-globin target itself (Audibert et al., 2002). The sickle cell transcript has no splicing aberrations and perhaps cis-splicing occurs so rapidly that the 5’PTM constructs are unable to compete effectively against it. We therefore attempted to improve our trans-splicing levels by targeting mutations that disrupt normal cis-splicing of exon 1 in β-globin.
Four different point mutations within β-globin’s intron 1 5’ splice site (IVS-1 5’ss) each result in a hematological disease called β-thalassemia. Unlike sickle cell anemia, β-thalassemia can be caused by various mutations in the β-globin gene and often the mutation is population specific. The four mutations we chose have all been documented in specific ethnic groups and each causes impairment of normal β-globin cis-splicing (table 3.1) (Treisman et al., 1983; Kazazian et al., 1984; Atweh et al., 1987; el-Hazmi et al., 1995; Colah et al., 2004; Darwish et al., 2005). Specifically, in the absence of normal IVS-1 5’ss usage, two different cryptic sites are substituted: E105 and E127. The former is found in β-globin exon 1 position 105 and the latter in exon 1 position 127 (Treisman et al., 1983; Atweh et al., 1987).

The two point mutations within the core endogenous 5’ss, IVS-1 +1 G->A and IVS-1 +1 G->T (hereafter referred to as “1A” and “1T”), produce the more severe clinical phenotype as all normal 5’ss usage is abolished (Treisman et al., 1983; Kazazian et al., 1984; el-Hazmi et al., 1995; Colah et al., 2004; Darwish et al., 2005). Since no normal β-chain peptides are produced, hemoglobin tetramers do not form and α-chains aggregate within the erythrocyte precursors, effectively destroying them and leading to the anemia that is a hallmark of β-thalassemia (reviewed in (Tuzmen & Schechter, 2001).

The mutations at position 5 of IVS-1, IVS-1 +5 G->C and IVS-1 +5 G->T (hereafter referred to as “5C” and “5T”), are not within the core 5’ss dinucleotide and therefore some normally spliced β-globin is produced (Treisman et al., 1983; Kazazian et al., 1984; Atweh et al., 1987; Colah et al., 2004). Since some free α-chains can now be
shuttled into hemoglobin tetramers, the anemia is less severe and these mutations are usually associated with the $\beta^+$ clinical phenotype. The 5C and 5T mutants also reflect the spectrum of clinical severity possible in $\beta$-thalassemia even when one mutation is used over another at the same position. The substitution of a cytidine at position 5 in intron 1 decreases the amount of normally spliced $\beta$-globin more so relative to when thymidine (uridine) exists at that site. This in turn affects the ratios of $\beta$ to $\alpha$ chains and leads to differences in clinical phenotype (table 3.1).
Table 3.1 The four β-thalassemia mutations used in this study.

Each mutation has been described in patients within the ethnic population specified. The third column describes both the molecular and clinical phenotype of each mutation. The numerical superscripts refer to the following studies (full citation provided in References section): ¹Treisman, R et al. 1983. ²el-Hazmi, MA et al. 1995. ³Darwish, HM et al. 2005. ⁴Colah, R et al. 2004. ⁵Atweh, G et al. 1987. ⁶Kazazian, H et al. 1984.

<table>
<thead>
<tr>
<th>B-thalassemia Mutation</th>
<th>Ethnic Group</th>
<th>Correctly Spliced B-globin mRNA Produced?</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVS1 +1 G-&gt;A</td>
<td>Jordanians, Egyptians, Syrian, Palestinians</td>
<td>No - cryptic only¹. B⁰-thal⁴</td>
</tr>
<tr>
<td>IVS1 +1 G-&gt;T</td>
<td>Asian Indians⁴,⁶</td>
<td>No - cryptic only⁶. B⁰-thal⁴</td>
</tr>
<tr>
<td>IVS1 +5 G-&gt;C</td>
<td>Asian Indians⁴,⁶</td>
<td>Some, but mostly cryptic⁶. Severe B⁺-thal⁶</td>
</tr>
<tr>
<td>IVS1 +5 G-&gt;T</td>
<td>Mediterranean, Northern European⁵</td>
<td>Some, but also cryptic⁶. B⁺-thal</td>
</tr>
</tbody>
</table>
We hypothesized that since cis-splicing is impaired in all four β-thalassemia mutations, our 5’PTM constructs would effectively compete against them for spliceosomal recognition. We tested this with co-transfection experiments in human embryonic kidney cells (HEK293T) using the same three 5’PTM constructs described in section 3.2 as well as the same β-globin backbone plasmid (simply replacing the sickle β-globin cassette with a β-thalassemia version). We then assayed for levels of normal splice product produced.

3.4.3 Results

Each 5’PTM construct (described in section 3.2) was co-transfected along with one of the four different genomic β-globin thalassemia expression vectors into 293T cells and trans-splicing was assessed 48 hours later using total RNA in a one-step RT-PCR reaction. Using trans-splicing specific primers, we detected the expected product of 410bp indicating that PTM-driven restoration of the normally spliced β-globin transcript using each of the three 5’PTMs with each of the four β-thalassemia mutations had occurred (figure 3.22). The identity of the trans-spliced products was verified by sequence analysis. The mix controls for these experiments were derived in the same way as in the sickle β-globin studies (section 3.2) and again were indicative of a true intracellular trans-splicing event.
Figure 3.22 The four β-thalassemia mutants are corrected via 5’PTM-driven trans-splicing.
Various βthal RNAs are repaired via 5’PTM trans-splicing in 293T cells. 5’PTM (5’PTM1, 5’PTMIDSE, & 5’PTMCodon) and genomic βthal-globin expression cassettes were co-transfected into 293T cells and total RNA was harvested for analysis by RT-PCR. Primers specific for trans-spliced products were used to amplify the repaired RNA and yielded a DNA product of 410bp. The products were purified from an agarose gel and sequenced to confirm their identity. A. Testing the β-thalassemia T mutants. 1T and 5T refer to β-thalassemia target mutants βthal +1G>T and βthal +5G>T respectively. The transfection performed is listed above each lane in the gel. Trans-splicing is expected to yield a product of 410bp. The molecular weight marker band is 400bp in all cases. B. Testing the β-thalassemia A and C mutants. 1A and 5C refer to β-thalassemia target mutants βthal +1G>A and βthal +5G>C respectively. Co-transfection experiments and RNA analysis were performed as described in part A.
Total RNA was then amplified via one-step RT-PCR using a primer set that bound the 5’UTR and exon 2 of β-globin. This experiment detected the effects of each β-thalassemia mutation alone as well as each 5’PTM’s ability to restore the normal splice product since the assay allowed for amplification of both cryptic and normally-spliced RNAs (figure 3.23). We found that the β-thalassemia point mutations produced splicing profiles similar to what had been previously described in the literature (figure 3.23) (Treisman et al., 1983; Atweh et al., 1987). The mutations within the core 5’ss at +1 (1A and 1T) produced no detectable 182bp normal splice product whereas mutations at position +5 (5C and 5T) produced some, but the cryptic products still predominated (figure 3.23, lanes 3-6). Co-transfection of 5’PTM1, 5’PTMDISE, or 5’PTMCodon expression cassettes with any of the four β-thalassemia mutants resulted in the appearance or increase (relative to cryptic products within the same lane) of the 182bp normal splice product (figure 3.23). For example, compare lane 3 (1A mutant alone) with lane 14 (1A mutant with 5’PTMDISE).
Figure 3.23 5’PTMs can restore the normal splice product in each β-thalassemia mutant.
The two cryptic 5’ splice sites, E127 and E105, produce RT-PCR products that differ in size from that generated via normal splicing (144 bp and 166bp versus 182bp). When either of the three 5’PTMs are used, the normal splice product appears (for β-thal 1A and 1C) or increases relative to the cryptic products in the same lane (for β-thal 5C and 5T). This is illustrated by comparing the band types produced by each β-thalassemia mutant alone to the 5’PTM-treated counterpart in the figure. For example, compare lane 3 (1A mutant alone) with lane 14 (1A mutant with 5’PTMDISE). The molecular weight marker used is the same for all marker lanes and the band sizes are noted next to lane 1 (125bp, 150bp, 175bp from bottom to top).
<table>
<thead>
<tr>
<th>Splicing</th>
<th>RT-PCR Product Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>182bp</td>
</tr>
<tr>
<td>Cryptic E127</td>
<td>166bp</td>
</tr>
<tr>
<td>Cryptic E105</td>
<td>144bp</td>
</tr>
</tbody>
</table>
We then used a HaeIII restriction enzyme assay to verify and quantify the levels of normally-spliced product. A HaeIII site exists between the locations of the cryptic splice sites and the normal 5’ splice site in β-globin exon 1 and thus only the normally-spliced RT-PCR product will retain this restriction site after splicing (figure 3.24). We used this unique HaeIII characteristic to verify the presence of the normal splice product by digesting some of the gel purified 182bp products seen in figure 3.23 and this result is shown in figure 3.24. The 133bp and 49bp HaeIII digestion fragments seen indicate the presence of the normally spliced product. However, a seemingly “182bp” band is retained in the 182bp digestion lanes. It is unlikely to be the result of incomplete digestion since it is not present in the βwt lane. A more likely possibility is that it is a heteroduplex formed between the normal 182nt and abundant cryptic 166nt cDNAs that is unable to be digested by HaeIII. Attempts to sequence this fragment suggested the presence of more than one sequence in the band, further supporting the heteroduplex hypothesis. We chose to include this heteroduplex in the following quantification analysis since it may have acted as a “sink” for the cDNAs.

As expected, the gel purified 166bp and 144bp bands were not cleaved by HaeIII. This illustrated the assay’s ability to distinguish between cryptic and normal dsDNA fragments (figure 3.24, lanes 14-17). It should also be noted that due to the close similarity among product sizes, some carryover was expected during excision of the individual gel slices. This is seen by the presence of 166bp and 144bp bands in the 182bp and 166bp gel extract lanes respectively. Since we extracted all band types in a lane as
one single gel slice prior to HaeIII digestion, carryover was not an issue during the actual quantification.
Figure 3.24 Testing a HaeIII restriction enzyme assay for potential use in quantifying β-thalassemia repair levels.

After RT-PCR using primers that bind both normal and cryptic splice transcripts, a HaeIII restriction site exists only in the normal splice product thus allowing for distinction between the possible outcomes. The three different RT-PCR products (normal, cryptic E127, cryptic E105) shown in figure 3.23 were excised individually from the agarose gel and purified. Each band was then digested overnight with HaeIII and run on an agarose gel as shown here. The two fragments resulting from HaeIII digestion of the normal 182bp product are 133bp and 49bp, both of which can be seen in the figure. The normal splice product inherently produced by the 5C and 5T mutants is also cleaved as expected. The two cryptic RT-PCR products, 166bp and 144bp, do not produce the 133bp and 49bp HaeIII digestion fragments. This illustrates the assay’s utility in quantifying repair levels via appearance of HaeIII digestion fragments (we used the 133bp band specifically). The “182bp” band remaining in some lanes is likely an indigestible heteroduplex product (explained in text). The 166bp band seen in the 182bp product digestion lanes is carryover from the gel slice excision process. The same is true for the 144bp band seen in the 166bp product digestion lanes.
We quantified the levels of normal splicing restored via each 5’PTM using one of the HaeIII digestion fragments (133bp) as the indicator for “normal” with this identity confirmed via sequencing of the fragment. Quantifying one of the HaeIII digestion products versus using the original 182bp band allowed us to account for the indigestible heteroduplex while providing a conservative, yet true, assessment of repair efficiency. We used Quantity One software (Bio-Rad) to quantify the intensity of the 133bp HaeIII fragment in each lane and compared this to the intensities of the other bands in the same lane (figure 3.25, representative gel image).

As shown in figure 3.25, we found that 5’PTMDISE and 5’PTM Codon can restore the normal splice product in the two β⁰-thalassemia mutations (1A and 1T) up to levels reaching 36% and these results are statistically significant (figure 3.25, p < 0.05 in all four cases). Both PTMs also significantly increase the baseline level of normal splice product in the 5C β-thalassemia case (figure 3.25, p < 0.05). The statistically significant decreases in cryptic products seen in the 1A, 1T and 5C mutants perhaps reflect a 5’PTM-driven increased conversion of pre-mRNA into normal rather than cryptic products (figure 3.25). The significant increases in heteroduplex product when comparing 5’PTMDISE and “No PTM” in the 1A and 1T cases may also be caused by the significant increase in normal product since we hypothesize that the heteroduplex is due to basepairing between normal and E127 cryptic cDNAs. The E127 cryptic product is still in excess over the normal product even when repair efficiencies are at their highest.
hence any increase in normal product may be redirected to forming the heteroduplex as well as the normal dsDNA.

5’PTM1 did produce the normal β-globin splice product (where there was none in the untreated sample) in the β0-thalassemia 1T case, but its level was not high enough to be statistically significant. This PTM’s effect on normal splice levels was similarly not statistically significant in β-thalassemia 5C. Finally, none of the 5’PTMs resulted in a clear increase of normal splice product in β-thalassemia 5T (figure 3.25). The quantification data is summarized in figure 3.26. The means for each condition are plotted in bar graph form and the six statistically significant increased normal splice product levels are starred with a maximum of 36% repair attained in the case of 5’PTM Codon repairing the β-thalassemia 1A mutant.
Figure 3.25 Quantification of repair levels engendered by each 5’PTM targeting the four β-thalassemia mutants.

After harvesting RNA from transfected 293T cells, performing RT-PCR, and running the products on a gel, the normal (182bp), cryptic E127 (166bp) and cryptic E105 (144bp) RT-PCR products were digested with HaeIII and run on a high resolution agarose gel containing ethidium bromide (see gel, this figure). The 133bp HaeIII digestion fragment was used as the indicator for the normal splice product. Quantification of each band (excluding the other HaeIII fragment (49bp)) within each lane was performed using Quantity One densitometric software from the Bio Rad company. The intensity values were normalized for band nucleotide length and each band type quantity was expressed as a fraction of the total of all bands quantified using a particular 5’PTM/β-thalassemia mutant combination (see Methods for details). The four β-thalassemia forms are graphed separately and the label within each x-axis parentheses denotes the band type quantified. Each experiment was done in triplicate or greater and the error bars shown here indicate standard error of the mean. Significance testing for each β-thalassemia form was performed using the untreated (i.e. the β-thalassemia mutant alone) value for a particular band type (i.e. normal, cryptic E127, cryptic E105, or undigested heteroduplex) and comparing it to its PTM-containing counterpart. The different p values are denoted as follows: * p < 0.05, ** p < 0.01, *** p < 0.001.
Band Types
1. Indigestible (Undigested Heteroduplex)
2. Cryptic E127
3. Cryptic E105
4. Normal Splice Product
(HaeIII 133bp fragment)
Figure 3.26 Percent repair of β-thalassemia mutants using various 5'PTM constructs.
The percent of each band type present for each 5'PTM/β-thalassemia combination is shown here along with the untreated controls (each β-thalassemia mutant alone). The stars denote statistically significant increases in levels of normal splice product when treatment is used (i.e. 5'PTM is co-transfected).
3.4.4 Discussion

We have shown that 5’PTM repair can be applied to β-thalassemia and that splicing-impaired targets may be suitable for SMaRT therapeutics. Using mutations in the β-globin gene intron 1, we found that 5’PTMs can trans-splice to repair defects causing some forms of β-thalassemia at the RNA level (figure 3.22).

We estimated that our trans-splicing levels in the sickle cell experiments were below 5% when assessed by RT-PCR followed by restriction enzyme analysis (section 3.2). However, repair efficiencies were as high as 36% when determined in a similar fashion but with a splicing impaired target, namely the β-thalassemia splice mutants (figure 3.25). These combined results suggest that perhaps either the PTM must out-compete the endogenous splice site for the splicing machinery or the target itself must be hindered in its splicing ability. The β-globin gene is normally spliced very efficiently (Audibert et al., 2002) and perhaps this is why, although we can detect the repaired transcript in the sickle cell case, it is not in high abundance. The sickle cell mutation does not impact the gene’s ability to splice, but instead leads to a structurally compromised β-chain within the hemoglobin tetramer (Ingram, 1956, 1959; Perutz, 1960). Our 5’PTM constructs may have been unable to compete effectively against this efficiently cis-spliced target.
However, when using β-thalassemia targets involved with splicing aberrations, 5’PTMs (5’PTMDISE and 5’PTMCodon specifically) significantly increased the levels of normal β-globin RNA product made versus that in untreated samples (figure 3.25). Perhaps the intron 1 β-thalassemia splice mutants cause a “stall” in the formation of the spliceosomal complex during which the 5’PTM can bind to the target and provide a viable alternative to the endogenous 5’ss.

Another possibility is that the suboptimal splice donor options in these β-thalassemia transcripts make the environment more amenable to basepairing with the 5’PTM binding domain. The β-thalassemia splice mutants dampen or, as in the +1 mutants, completely silence the naturally selected splice donor. If the U1 snRNP is thus presented with solely mediocre 5’ splice site options, it may only weakly interact with the pre-mRNA (Freund et al., 2003). Perhaps this “vulnerability” is exploited by the 5’PTM’s binding domain as its exact basepairing may provide the more stable interaction. After PTM binding, the splicing factors may then recognize the 5’PTM’s splice donor sequence as the strongest option in the region and this may lead to increased \textit{trans}-splicing levels.

A more general mechanism might be that the splicing impairment in these target transcripts causes them to remain in the nucleus for a longer period of time and hence continue to be a viable option for SMaRT. Work by Custodio and colleagues has shown that splicing aberrations in β-globin intron 2 leads to retention of pre-mRNA transcripts.
in the nucleus and hampered exon junction complex formation (Custodio et al., 1999; Custodio et al., 2004).

A final possibility for the apparent increased effectiveness of 5’PTMs in the β-thalassemia versus sickle cell anemia experiments involves the selective destabilization of certain transcripts and their impact on the repair calculation. Perhaps the β-thalassemia cryptic transcripts are undergoing nonsense mediated decay (NMD) and hence the net level of *cis*-spliced RNA is lower here than in the sickle cell anemia case. This would decrease the denominator in the β-thalassemia repair efficiency ratio: [normal splice product]/[heteroduplex + cryptic E127 + cryptic E105 + normal splice products] and any *trans*-splicing would have a greater apparent effect compared to identical levels of *trans*-splicing in the sickle cell anemia example. Analysis of the mRNA sequences produced by the two cryptic 5’ splice sites reveal stop codons located further than 50-55nt upstream of the nearby splice junction and thus these cryptic transcripts may be susceptible to NMD (Maquat, 2005). The cryptic E105 site results in a downstream premature stop codon that is 184nt upstream of the exon 2/exon 3 splice junction. The E127 site leads to a downstream premature stop codon that is 131nt upstream of the same junction. A 2002 study by Danckwardt and colleagues examined the possibility of NMD of cryptic transcripts produced from a β-thalassemia 5A mutation that are identical to those described here (Danckwardt et al., 2002). They found that the E127 cryptic RNA was subject to NMD. This study bolsters our suggestion that perhaps 5’PTM repair is more detectable in our β-thalassemia examples than in the sickle cell case since the *trans-*
spliced product potentially holds a selective advantage in stability in the former. However, Danckwardt and colleagues also found that the E105 transcript was not susceptible to NMD and in fact, its usage as a cryptic 5’ss was dependent on E127 cryptic activity. This conclusion casts doubt on E105 transcripts skewing repair efficiency values in the β-thalassemia case, but it also poses an intriguing possibility for 5’PTM design. Perhaps one might extend the PTM’s binding domain so that it basepairs with the E127 cryptic region and in doing so, prevents both cryptic sites (E105 and E127) from being utilized. The spliceosome might be even more apt to choose the 5’ss provided by the 5’PTM.

This work has shown that SMaRT can repair certain β-thalassemic transcripts and that its applicability may extend to other disease targets with abnormal splicing as an underlying pathology. A pre-mRNA’s endogenous cis-splicing rate and the resulting transcript’s stability play important roles in the success of trans-splicing therapy and hence should be considered when selecting targets for SMaRT. Our diverse 5’PTM designs also appear to influence the repair levels achieved since, among other differences, unique 5’ss were used in each construct and the PTMs repaired the targets to varying degrees. We are currently designing a 5’PTM identical to 5’PTMDISE save for the DISE sequence itself in hopes of revealing how this particular design characteristic may impact repair efficiencies. Our 5’PTMs could also include other ISEs like the GGG motif used in α-globin introns or the 4.1R putative Fox-2 binding sequence, both of which are used during erythropoiesis (McCullough & Berget, 2000; Ponthier et al., 2006)
3.4.5 Materials and Methods

Plasmids

The 5’PTM expression plasmids were constructed and used as described in the Methods portion of section 3.2. The β-thalassemia β-globin genomic sequences were cloned into the RNA polymerase II expression vector pcDNA3.1(+)Hyg (Invitrogen) using NheI and BamHI restriction enzymes (NEB). Four different β-thalassemia IVS-1 point mutations were used: +1 (G to A), +1 (G to T), +5 (G to C), and +5 (G to T). All four mutations were characterized via RT-PCR plus agarose gel electrophoresis (Superscript III/ Platinum Taq One-Step RT-PCR by Invitrogen) and shown to activate two different cryptic 5’ splice sites (E105 and E127) leading to shortened amplimers.

Transfections, RNA Isolation, and RT-PCR

As was described in section 3.2.

HaeIII Assay

Total RNA from β-thalassemia/5’PTM co-transfection experiments was isolated as previously described and amplified via one step RT-PCR using primers that would amplify both cryptic and normally spliced transcripts (primers bound 5’UTR and exon 2 of β-globin) using the following conditions: 56°C for 30 minutes, 94°C for 2 minutes, 25 cycles of (94°C 30 seconds/55°C 30 seconds/72°C 15 seconds) then 7 minutes at 72°C with a final hold at 4°C. RT-PCR products resulting from both cryptic and normal splicing were isolated as single gel slices (one per lane) and purified. The purified products were then digested overnight with HaeIII (NEB). The fragments were resolved
on a 3.5% agarose-1000 gel (Invitrogen) containing ethidium bromide and the intensities of normal (133bp HaeIII fragment) and abnormal (cryptic 166bp/144bp fragments or indigestible heteroduplex fragment present after complete HaeIII digestion) splice products were quantified using Quantity One densitometry software (Bio-Rad).
Experiments were replicated 3-5 times and band density (intensity) was normalized for fragment nucleotide length. The means of the replicates were used to calculate relative amounts of each band type (per PTM/β-thal mutant combination) and graphed in figure 3.25. One of these ratios was used in calculating the repair efficiency: [normal splice product]/[heteroduplex + cryptic E127 + cryptic E105 + normal splice products].
Statistical significance testing was performed on band length normalized relative intensity data using Bonferroni pairwise comparison testing with a 95% confidence interval using GraphPad Prism (GraphPad Software Inc.). Each band type/PTM/β-thal mutant combination was tested against the comparable “No PTM” combination to determine if there was a significant difference in relative mean intensity.
Chapter 4: Conclusions and Future Work

4.1 DNA Repair Using the L. lactis Group II Intron

4.1.1 The L. lactis Group II Intron’s Target Specificity Results in a Potentially Safer DNA Repair Therapeutic

The utility of any potential clinical therapy distills down to two questions. Is it safe and is it effective? We have shown here that the Ll.LtrB targets its DNA insertion site with high specificity in an *E. coli* model and we are beginning work to assess this using a β-globin target. The mobile group II intron provides not only specificity, but also flexibility in target selection and our lab has previously illustrated the retargeting capabilities of Ll.LtrB to therapeutically relevant DNA targets.

The challenge for its therapeutic development now lies in its ability to function in mammalian cells. Reconstituted RNPs have been shown to mobilize into plasmid targets in mammalian cells, but we have been unable to generate RNPs within the cells themselves. Perhaps this is not a barrier to some therapeutic applications though. The utility of *ex vivo* gene therapy with subsequent bone marrow transplant for certain hematological conditions may warrant the use of reconstituted RNPs. The *ex vivo* RNP delivery would provide the benefits of targeted genomic insertion without the need for intra-cellular assembly of the RNP. The mere presence of the RNP in the mammalian cell may not be sufficient though, as will be discussed in the next section.
4.1.2 Heterochromatin and the Mammalian Cell Environment May Pose

Challenges to This Therapeutic Approach

A functional RNP, even if successfully delivered to the nucleus via a nuclear localization signal, may still encounter difficulty when attempting to integrate into histone-laden chromosomal DNA. Targeting actively transcribed regions is one way to circumvent some of this potential issue, but that approach alone is not a panacea. Care must be taken to avoid disrupting expression of surrounding genes and, as in lentiviral therapeutic applications, position effects may squelch long-term expression of the intron’s domain IV transgene.

Furthermore, unlike lentiviruses that have had years of evolutionary selection to “learn” how to successfully navigate through mammalian cell hosts, the Ll.LtrB group II intron has evolved to thrive within its native bacteria. The foreign mammalian cell and its chromosomal DNA may pose unseen barriers to Ll.LtrB retrohoming. One might invoke the RNA World hypothesis to suggest that since mobile group II introns may be ancestors to transposons and spliceosome machinery, the jump to mammalian cells might prove easier than it appears (Gesteland et al., 1999). The hypothesis proposes that RNPs established a mutual symbiotic relationship with eukaryotes to become ancestors of modern-day cellular components (Gesteland et al., 1999). This shared history is no guarantee for safe passage though. The mammalian cell has developed robust defense mechanisms against foreign invaders throughout its evolution (for example, anti-viral double-stranded RNA silencing processes) and we cannot predict how these pathways
will affect the RNP. Unlike lentiviruses, that have co-evolved to evade mammalian host defenses, the RNP is relatively unprepared for what the mammalian cell may unleash upon it. Perhaps the cellular environment will prove quite harsh for RNPs traveling to the nucleus and attempting to integrate into the chromosomal DNA. The potential impact of the host cell itself is thus an important consideration as we strive to bring this unique therapeutic modality to fruition.

4.2 RNA Repair Using 5’PTMs and SMaRT Technology

4.2.1 The Potential Impact of Nuclear Double-Stranded RNA Processes on SMaRT

As was touched on in the previous section, therapies utilizing double-stranded RNA intermediates may be targets for nuclear RNA silencing. These silencing processes could involve microRNAs (via Drosha/DGCR8), adenosine deaminases (ADAR), or nuclear interferon responses.

MicroRNAs (miRNAs) are small single-stranded RNAs that depress gene expression by binding to a target mRNA’s 3’UTR and inhibiting translation (Olsen & Ambros, 1999). Unlike siRNA activity that results in mRNA target degradation, miRNAs exert their repressive effect while usually sparing the target transcript (Olsen & Ambros, 1999). It is thought that miRNAs exist as important regulators of gene expression in normal cellular processes (Carrington & Ambros, 2003; Lewis et al., 2005) and the
potential inappropriate recognition of double-stranded RNA therapeutics (like the PTM’s binding domain/target pre-mRNA complex or double-stranded structures resulting from PTM intra-molecular basepairing) by this pathway might stymie any therapeutic effect. It is therefore important to weigh this possibility using our current understanding of miRNA genesis.

The natural process begins with long RNA transcripts looping to form large double-stranded RNAs called pri-miRNAs that are then cleaved to smaller 65nt hairpins called pre-miRNAs by the coordinated efforts of two proteins, Drosha and DGCR8 (Lee et al., 2002; Lee et al., 2003; Denli et al., 2004; Gregory et al., 2004; Han et al., 2004). These pre-miRNAs are then exported out of the nucleus to be further processed by DICER and one strand joins the RISC RNP as the final miRNA (Hutvagner et al., 2001; Hutvagner & Zamore, 2002; Yi et al., 2003; Hutvagner, 2005).

The requirements for processing via Drosha/DGCR8 in the nucleus are not fully understood, but some questions have been answered. Basepairing near the end of the pri-miRNA stem and the presence of unstructured ssRNA flanking sequences both appear to be vital characteristics (Zeng et al., 2003; Zeng & Cullen, 2005; Han et al., 2006). Perhaps one of the most important elements for our purposes is whether the loop in the pri-miRNA is critical for Drosha/DGCR8 processing. Looking at this in another way, can Drosha/DGCR8 process dsRNA resulting from the binding of two separate RNAs? This would be the case for the PTM binding domain/pre-mRNA target complex and hence the answer would address one of our core concerns. Work by Han and colleagues suggests
that the loop is unnecessary, but these experiments were performed in vitro and Drosha/DGCR8 may function differently within the cell itself (Han et al., 2006). Furthermore, other investigators have found that larger rather than smaller loops enhance processing activity in human cells. Although they did not test the “no loop” condition, their findings seem to suggest that the loop is important in cells (Zeng et al., 2005). The required single stranded flanking sequence length also appears to differ in in vitro versus in vivo experiments. Studies have found that Drosha/DGCR8 proteins are able to process pri-miRNAs with only 10nt ssRNA segments in vitro whereas longer segments seem to be required within cells (Chen et al., 2004; Zeng & Cullen, 2005).

Experimental testing should reveal if Drosha/DGCR8 acts on PTM/target complexes containing long stretches of double-stranded RNA flanked by single stranded segments. An artificial PTM/pre-mRNA construct could be expressed via RNA polymerase II (using an approach similar to what is described in (Tran et al., 2004)) and tested for Drosha processing within cells. Perhaps an initial experiment might simply entail using an in vitro transcribed and annealed PTM/pre-mRNA long double-stranded RNA complex in an in vitro Drosha assay. A schematic of a potential substrate for this experiment is shown in figure 4.1. Follow-up experiments within cells would be necessary to confirm any in vitro findings though.

Since the PTM may likely form intra-molecular bonds, especially between complementary binding domains and cargo sequences or when using a 3’PTM safety, a hairpin structure may form that could also be susceptible to Drosha/DGCR8 processing.
Testing of potential PTM hairpins in Drosha assays is thus also important since Drosha/DGCR8 activity could deplete therapeutic substrate levels.
Figure 4.1 Hypothetical PTM/target pre-mRNA dsRNA for use in Drosha activity testing.
The hypothetical construct has both a target pre-mRNA strand in yellow and a PTM RNA strand in blue (a 5’PTM in this case). This double-stranded RNA could be generated either via *in vitro* transcription/annealing or intra-cellularly using an RNA polymerase II expression system. Potential processing via Drosha could also be assayed either *in vitro* or *in vivo*, but *in vitro* results would require *in vivo* confirmation.
ADAR or “Adenosine Deaminases Acting on RNA” is a modification pathway whereby certain adenosines in long dsRNAs are converted to inosines via deamination (Bass, 1995, 2002). ADAR is thus a potential source for codon variation in genes (ADAR reviewed in (Nishikura, 2006)). Recently, it has been shown that pri-miRNAs can be targets for this pathway and that the pathway may not only shuttle substrate away from Drosha, but it may also use the nuclease TudorSN to degrade these substrates (Yang et al., 2006). If the PTM/target pre-mRNA or PTM hairpin structures prove susceptible to Drosha activity, we must then also consider potential effects this alternative pathway may have on our therapeutic approach.

Finally, we must consider cellular interferon responses and their potential impact on our PTMs. Cellular interferon responses include PKR and 2’-5’OAS (reviews in (Justesen et al., 2000; Garcia et al., 2006)). PKR recognizes cytoplasmic double-stranded RNA and although its activity can lead to nuclear effects like transcriptional activation of genes via an activated NFkB transcription factor, this stems from the initial cytoplasmic dsRNA detection (Garcia et al., 2006). The dsRNA interaction between PTM and pre-mRNA target forms within the nucleus and therefore it is unlikely to stimulate PKR activity. Some related work done using nuclear expression of long dsRNAs for RNAi applications appears to support this notion. Tran and colleagues did not detect any PKR activity after RNA polymerase II-directed nuclear expression of long dsRNAs for downstream use in RNAi (Tran et al., 2004). Strat et al. did not detect an intracellular
interferon response after RNA polymerase III nuclear expression of long dsRNAs (Strat et al., 2006). Since 2’-5’ OAS is also primarily considered a sentry for cytoplasmic dsRNA, we do not anticipate activating its activity with our nuclear-based dsRNA (Justesen et al., 2000).

4.2.2 Target Choice and SMaRT

Our attempts to quantify repair efficiencies in the sickle cell experiments and subsequently using the β-thalassemia splice mutants revealed much about the nature of targets accessible to SMaRT. The improvement in trans-splicing efficiency when targeting β-thalassemia rather than sickle cell anemia suggests that splicing-impaired pre-mRNA targets or mutations that produce deleterious transcripts (like the cryptic cis-spliced products with their premature stop codons marking them for nonsense mediated decay) are particularly receptive to SMaRT RNA therapy.

There are a number of clinical conditions that may be good candidates for SMaRT. For example, a common lysosomal storage disorder called Gaucher’s disease is an autosomal recessive condition caused by mutations in the GBA gene that leads to insufficient levels of its gene product. The absence of this enzyme causes build-up of glucosylceramide in certain cell types that can have orthopedic, neurologic, or cardiac consequences depending on the form of the disease (Goldblatt, 1988). Beutler and colleagues (Beutler et al., 1992; Beutler, 2006) reported on three common mutations in the Ashkenazi population including an IVS-2 point mutant in its 5’ splice site. Employing
a 5’PTM therapeutic approach here may be effective due to the underlying splicing pathology. The 5’PTM might also prove to be an efficient and specific treatment modality since only the first two of 11 total GBA exons would require replacement and the binding domain would target the therapeutic to this transcript specifically.

A second candidate disease might be phenylketonuria or PKU. This condition is caused by a deficiency in phenylalanine hydroxylase and it can lead to severe mental retardation if left untreated. Various mutations in the PAH gene leading to PKU have been identified, but two fairly common variants may be particularly good targets for SMaRT. DiLella and colleagues described one of the first PKU-causing mutations in 1986. It was a G to A point mutation in the 5’ splice site of intron 12 (DiLella et al., 1986). Dworniczak et al described an intron 10 mutation that was common in parts of Europe and it too caused a splicing abnormality (Dworniczak et al., 1991). The mutation activated a cryptic splice site in the intron and caused improper splicing in that region. Either of these mutations might be clinically relevant and attainable targets for PTM trans-splicing RNA repair.

SMaRT technology possesses much promise as an RNA repair mechanism and as we learn more about the spliceosome, alternative splicing, and the molecular pathology of various conditions, we will hopefully harness its full potential for treating human disease.
References


Callard RE, Armitage RJ, Fanslow WC, Spriggs MK. 1993. CD40 ligand and its role in X-linked hyper-IgM syndrome. Immunology today 14:559-564.


Biography

Monique Natasha Kierlin-Duncan


Education:

Undergraduate honors thesis topic: Interaction Between the N-Terminus of the Neurofibromatosis 2 Protein (Merlin) and Human Gene Products.
Advisor: David R. Cox M.D./Ph.D.

Masters thesis: Autonomy in Genetic Counseling.


Publications:

(*co-first author publication)