NONSENSE MEDIATED DECAY IN BETA-THALASSEMIA

LOO POOI ENG

(BSc. Hons)

A THESIS SUBMITTED FOR THE DEGREE OF MASTER

OF SCIENCE

DEPARTMENT OF PAEDIATRICS

NATIONAL UNIVERSITY OF SINGAPORE

2009

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor, A/P Samuel Chong for his insightful advice and guidance throughout the course of my studies. Thank you for sharing your expertise and I have truly learnt a lot from you.

A very special thank to Dr. Wang Wen for being a great mentor. Though you are not in the lab anymore, your unwavering help and guidance is deeply appreciated.

To Arnold and Dr. Felicia Cheah, thank you for sharing with me your laboratory expertise and making sure things run smoothly in the lab.

I would also like to express my appreciation to all my lab members and colleagues. Thank you for the encouragement and support. I really appreciate the help rendered and it was a great pleasure working with all of you.

Last but not least, I am very grateful to my family and Chia Yee for their endless love, support and encouragement.

Thank you.

Table of contents

Abstract of thesis

β-thalassemia is one of the most common single gene disorders that results from reduced or non-production of β-globin chains. Most of the β-thalassemias are caused by frameshift or nonsense mutations within the β-globin gene or its immediate flanking sequences, which produce premature termination codons (PTCs). In β-thalassemia, nonsense mediated decay (NMD) offers protection from the potentially deleterious effects of PTCs, which will result in the synthesis of truncated proteins. Hence, the relationship between NMD with the location of PTCs in β-thalassemia was studied by analyzing the gene expression levels of codon 41/42 mutations (-TTCT) (located at exon 2), codon 121 mutations (G→T) (located at exon 3) and IVS 2-654 mutations (C→T) (located at intron 2). Steady state gene expression study in HeLa cells showed to comply with the 50-55 nucleotide boundary rule of NMD. The observation was confirmed by using a transcriptional pulse-chase strategy, HeLa Tet-Off system. Gene expression study in MEL (mouse erythroleukemia) cells showed similar results except for IVS 2-654 mutation, whereby a relatively low gene expression level was observed. This interesting observation suggests that the 50-55 nucleotide boundary rule does not apply to IVS 2-654 and we postulated that the extra 73 nucleotides intronic sequences in IVS 2-654 mutant transcript is able to trigger NMD and subsequently contributed to the recessive inheritance phenotype of the mutation.

Summary

β-thalassemia is one of the most common monogenetic disorders in human, characterized by the reduced or non-production of β-globin chains. To date, there are almost 200 mutations that have been identified and most homozygotes suffer from a severe syndrome and require regular blood transfusions to survive. These mutations normally result in aberrant transcripts with premature termination codons (PTCs) which will lead to the synthesis of truncated proteins. Hence, there is a post-transcriptional mechanism to control the quality of mRNA function by selectively degrading mRNAs that prematurely terminate translation due to the harboring of PTCs. A general rule has been postulated for the identification of PTCs that triggers nonsense mediated decay (NMD) in transcripts: if PTCs are located more than 50-55 nucleotides upstream of the 3' most exon-exon junction, the mRNA will be subjected to NMD, however, PTCs located downstream of this region will not be targeted for NMD.

In this study, the relationship between NMD and the location of PTCs in β-thalassemia was studied by looking at gene expression level of codon 41/42 mutations (-TTCT) (located at exon 2), codon 121 mutations (G→T) (located at exon 3) and IVS 2-654 mutations (located at intron 2) (C→T). Steady state gene expression levels of these three mutations in HeLa cells, a non-erythroid cell lines, has shown to comply with the 50-55 nucleotides boundary rule. PTC at codon 60/61 for the codon 41/42 mutation was observed to trigger NMD and averaged about 50% of wild type's gene expression level. Both codon 121 and IVS 2-654 mutations that created a PTC at codon 121 managed to escape from NMD and hence, averaged a high gene expression level, which are about 100-110% of wild type expression levels. This

steady state gene expression levels in HeLa cells were confirmed by using HeLa Tet-Off system, a transcriptional pulse-chase strategy.

In addition, mouse erythroleukemia cells (MEL) stable cell lines, carrying either the wild type of mutant human β-globin genes were developed in order to study the actual gene expression level of the three mutations in the erythroid environment. Erythroid differentiation was induced in these stable cell lines with the addition of DMSO to replicate the *in vivo* matured erythroblast stage. As expected, codon 41/42 mutation was subjected to NMD and averaged only 11% of wild type expression level. For codon 121, although this mutation averaged a slightly lower gene expression level than expected, a substantial amount of mutant transcripts was still detected. Interestingly, IVS 2-654 mutation averaged a relatively low expression level, which was only 15% of wild type transcript level. This finding suggested that the 50-55 nucleotides boundary rule does not apply to IVS 2-654 mutation, which is located in intron 2. Hence, we hypothesized that the extra 73 nucleotides intronic sequence of IVS 2-654 mutant transcript plays a crucial role in eliciting NMD and subsequently contributes to the observed recessive inheritance phenotype in patients.

List of Tables

List of Figures

1.0 Introduction

1.1 Hemoglobin

Hemoglobin is an iron-carrying protein that can be found in red blood cells (RBC). The main function of hemoglobin is to bind oxygen and transport them from the lungs to all body tissues, and followed by transporting carbon dioxide from body tissues back to the lungs. Hemoglobin has four oxygen binding sites and each of the sites is formed by a polypeptide chain (globin) and a prosthetic group (heme). This specific structure hence forms the functional tetramer molecule of hemoglobin.

The globin chain of hemoglobin is made up of four protein chains that are arranged in matching pairs. These chains can be categorized into two families, namely the alpha globin and beta globin families. The alpha globin family comprises of the alpha chains (α) and zeta chains (ζ). As for the beta globin family, it includes the beta chains (β), A-gamma chains (Aγ), G-gamma chains (Gγ), delta chains (δ) and epsilon chains (ε). The β-gene cluster was found to be located on chromosome 11 and the α-gene cluster on chromosome 16 (Deisseroth, Velez et al. 1976; Deisseroth, Nienhuis et al. 1977; Deisseroth, Nienhuis et al. 1978). Further studies using in situ hybridization and gene mapping have placed β-gene cluster distal to band p14 on the short arm of chromosome 11 (Gusella, Varsanyi-Breiner et al. 1979; Jeffreys, Craig et al. 1979; Lebo, Carrano et al. 1979; Sanders-Haigh, Anderson et al. 1980) and the α-gene cluster was shown to be specifically located in band 16p13.3 at the tip of chromosome 16 (Koeffler, Sparkes et al. 1981; Nicholls, Jonasson et al. 1987).

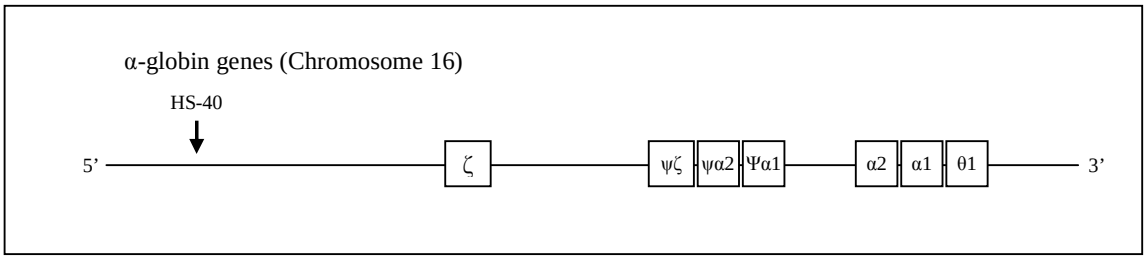
All these six different types of globin chains are found in normal human hemoglobin and expressed in a tissue- and developmental- specific fashion. The different requirements of oxygen during embryonic, fetal and adult life give rise to the synthesis of the different types of hemoglobins, specifically at different developmental stages (Higgs and Weatherall 2009). Basically, the structure of normal hemoglobins is formed by tetramers of two pairs of different globin chains. In normal adults, HbA ($\alpha_2\beta_2$) is the major component in total hemoglobin, which comprises of 97%, with HbA_2 ($\alpha_2\delta_2$) being the minor component, comprising 2.5%. In fetus stage, HbF ($\alpha_2\gamma_2$) is widely present and there are three embryonic hemoglobins in the embryo before the eighth week, the Hb Gower 1 ($\zeta_2\varepsilon_2$), Hb Gower 2 ($\alpha_2\varepsilon_2$) and the Hb Portland ($\zeta_2\gamma_2$).

During development, the embryonic to fetal globin switching will occur around the start of the 5^{th} week of gestation. The hematopoiesis process will take place at yolk sac islands and subsequently switch to fetal liver. This fetal hemoglobin production is continued up to shortly before birth. From twentieth week of gestation, hematopoiesis occurs in the spleen and the bone marrow, causing the fetal to adult hemoglobin switch near the perinatal period, when Hb F is replaced by Hbs A and A_2 over the first year of life (Weatherall and Clegg 2001).

1.2 α-globin gene

Human α-globin gene cluster is located at the telomeric region of the short arm of chromosome 16 (16p13.3), which occupies a region of around 70 kb (Weatherall and Clegg 2001). The functional α-globin genes are arranged in the specific order of 5' ζ_2-$\psi\zeta_1$-$\psi\alpha_2$-$\psi\alpha_1$-α_2-α_1-θ-3' (Figure 1) (Weatherall and Clegg 2001), surrounded by widely expressed genes such as 3-methyladenine-DNA glycosylase gene (MPG), chromosome 16 open reading frame 35 (C16orf35) and lethal unless nuclear cap-binding complex (Luc7L) (Higgs and Weatherall 2009).

Figure 1. α-globin gene cluster on chromosome 16



In the α-globin gene cluster, there are three functional α-like genes present, namely the ζ, α_2 and α_1 genes. As previously mentioned, the ζ-globin gene is majorly found in embryos, forming the Hb Gower 1 ($\zeta_2\varepsilon_2$), Hb Gower 2 ($\alpha_2\varepsilon_2$) and the Hb Portland ($\zeta_2\gamma_2$). On the other hand, both the α_1 and α_2 genes are widely found in adult genes and are highly homologous as both genes contain 3 exons separated by two segments of intervening sequences (introns). The only differences between the α_1 and α_2 genes are the sequence within the intervening sequence 2 (IVS 2) and the 3' non-coding region. Interestingly, although these two genes encode identical proteins, the ratio of gene expression between α_2 to α_1 is approximately 3:1 throughout the whole developmental stage (Weatherall, Stamatoyannopoulous et al. 2001).

In addition to the functional α-like genes, there are three pseudogenes known as the $\psi\zeta_1$, $\psi\alpha_2$ and $\psi\alpha_1$ genes located on the α-gene cluster (Hardison, Sawada et al. 1986). These three pseudogenes are thought to be the relics of past revolutionary changes within the globin gene cluster. Also located on the α-globin gene cluster is the θ-globin gene that is situated at the 3' end. The conserved structure of this gene and the fact that it contains no inactivating mutations, suggested the possibility that it might be a functional gene. However, no protein has been found to be encoded by this gene thus far, which may be the reason why the actual role of the θ gene remains unknown (Weatherall and Clegg 2001).

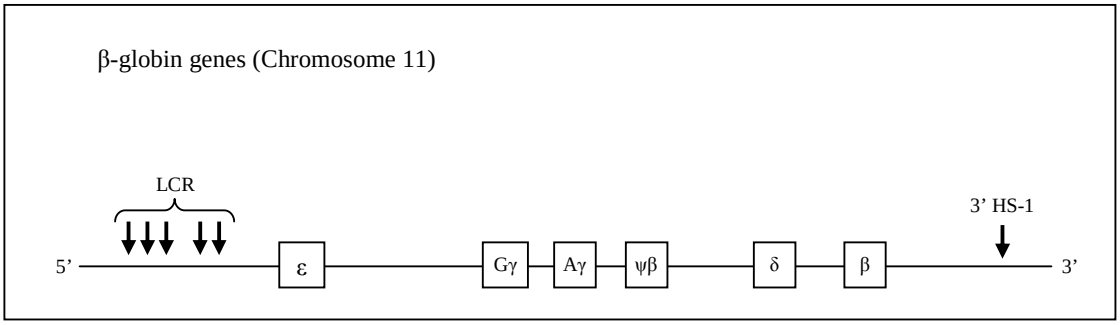
At the 5' end of the α-gene cluster, upstream of the ζ-globin mRNA CAP site, four highly conserved non-coding sequences, or multispecies conserved sequences (MCS) known as MCS R1-R4 were identified. These conserved sequences are thought to be involved in the regulation of α-globin gene expression. Of these four MCS, only MCS-R2 (also known as HS-40), located 40 kb upstream of ζ-globin mRNA cap site), has shown an essential effect on α-globin gene expression (Higgs and Weatherall 2009). This regulatory sequence was later termed α-MRE (major regulatory element) (Ribeiro and Sonati 2008).

From the DNA sequence of the α-globin gene cluster, the location of the α-globin locus lies very close to the telomere of the short arm of chromosome 16. The region is very G-C rich and contains many Alu family repeats. The linked α-globin genes are located within two highly homologous segments, spanning about 4-kb long. These regions are divided into homologous subsegments (termed X, Y and Z) by non-homologous elements (I, II and III) (Weatherall and Clegg 2001).

1.3 β-globin gene

Human β-globin gene cluster is located on the short arm of chromosome 11 (11p15.5). It includes the embryonic ε-globin gene, fetal globin genes $^{G}\gamma$ and $^{A}\gamma$, pseudogene ψβ, and the adult δ- and β-globin genes (Fritsch, Lawn et al. 1980). These functional genes are specifically arranged in the order of 5' ε-$^{G}\gamma$-$^{A}\gamma$-ψβ-δ-β 3' (Figure 2) and are expressed accordingly in the same order during development (Stamatoyannopoulos et al. 1974).

Figure 2. β-globin gene cluster on chromosome 11



Similar to α-globin gene, the β-globin gene cluster also contain an upstream regulatory element, located at the region before ε-globin mRNA CAP site. This regulatory element, termed locus control region (LCR), is physically comprised of five DNAse I-hypersensitive sites (HSs), spanning about 15 kb. This LCR element is found to be essential in regulating the expression of all the genes in the complex in erythroid tissue (Weatherall and Clegg 2001).

The β-globin gene cluster contains microsatellite repeats of $(CA)_n$ (usually of 17 dinucleotides repeats) and also a stretch of $(ATTTT)_n$ repeat between the δ and β gene. Moreover, numerous single nucleotide polymorphisms (SNPs) have been

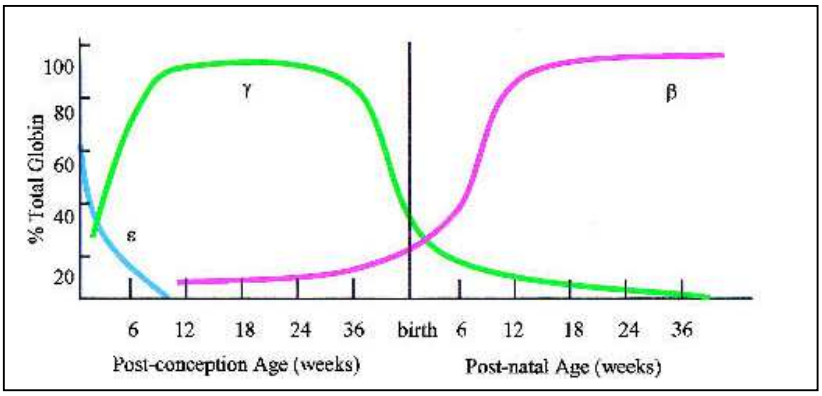
reported to be located on the β-globin gene cluster. Most of these SNPs are found to affect the cleavage sites for restriction endonucleases, hence creating restriction fragment length polymorphisms (RFLPs) (Weatherall and Clegg 2001).

Basically, β-globin gene spans about 1600 bp and encodes 146 amino acids. This stretch of 1.6 kb sequence is divided into three exons, intervened by two introns in between. These two introns are also termed as intervening sequences (IVSs). The β-globin gene contains a promoter that includes three positive cis-acting elements: TATA box, CCAAT box and duplicated CACCC motifs. These elements are located at 28 to 105 bp upstream of the mRNA cap site of β-globin gene. In addition, there is an enhancer that is found in intron 2 and 3' of the β-globin gene, which spans about 600-900 bp downstream of the poly (A) site (Thein 1998).

As previously mentioned, hemoglobin genes are expressed in a tissue- and developmental-specific fashion. For the β-globin gene cluster, there are two switches of gene expression of different β-globin genes at different sites where hematopoiesis occurs. During the first six weeks of gestation, the ε-globin gene is first expressed at a time when the γ- and β-globin genes are silent. This expression takes place in primitive, nucleated erythroid cells of the yolk sac. When the first switch takes place, the $^{G}\gamma$- and $^{A}\gamma$-globin gene expressions are activated in the fetal liver. Upon the activation of these γ-globin genes, the ε-globin gene will in turn be silenced. The expression of both $^{G}\gamma$- and $^{A}\gamma$-globin genes will be ongoing until birth where the second switch of gene expression occurs. Shortly after birth, expression of the β-globin gene and, to a smaller extent, the δ-globin genes are activated in the bone

marrow and spleen. Similarly, once the adult β-globin gene is activated, the γ-globin genes are in turn silenced (Figure 3) (Harju, McQueen et al. 2002).

Figure 3. Globin gene switching at different time point (Schechter 2008).



1.4 Thalassemia

Hemoglobinopathies are the most common human monogenetic diseases, characterized by the presence of abnormal hemoglobin in the blood. Approximately 7% of the world's populations are carriers of different inherited disorders of hemoglobin, with 60% of total and 70% pathological being in Africa. Generally, hemoglobinopathy can be classified into three major classes: (1) Structural variants in hemoglobins when the mutation alters the amino acid sequence of a globin chain, changing the physical properties and producing the clinical abnormalities; (2) Thalassemias arise when there is a reduction in the amount of normal hemoglobins that are being produced and; (3) a more diverse group of conditions which is characterized by the synthesis of high levels of fetal hemoglobin in adult life, commonly known as hereditary persistence of fetal hemoglobin (HPFH) (Weatherall 2001).

Thalassemia was initially known as Cooley's anemia as it was first identified by Thomas Cooley and Pear Lee in 1925 in four young children with anemia and splenomegaly, enlargement of the liver, discoloration of the skin and of the sclera and no bile in the urine (Eleftherious 2003). Following that, many Cooley's anemia cases were reported and it was observed to be predominant in Mediterranean races. Thus, this disorder was later named thalassemia, from the Geek word 'θαλασσα', which means "the sea", by Whipple and Bradford in 1932, in order to associate the disease with the Mediterranean area (Weatherall and Clegg 2001).

After 1940, the genetic nature of thalassemia became clear and it has been found that thalassemia occurs not only in the Mediterranean region but also in the Middle East, the Indian subcontinent and Southeast Asia. The similarity in the distribution of thalassemia and the areas in which malaria was endemic indicates that the positive selection for maintaining the high frequency of thalassemia might be malaria. It is also thought that the small genetic adjustment caused red blood cells to prevent parasite to survive and multiply. Thus thalassemia carriers were more likely to be able to survive malaria as compared to healthy individuals and this has contributed to the increasing number of carriers over the years (Weatherall and Clegg 2001; Eleftherious 2003).

By the early 1970s, it became apparent that there are many forms of thalassemias. All forms of thalassemia are characterized by the absence or reduced output of one or more of the globin gene chains of hemoglobin, leading to the imbalanced globin chain synthesis (Weatherall, Stamatoyannopoulous et al. 2001). This is the hallmark of all the thalassemia syndromes, which subsequently causing the

ineffective erythropoiesis and short life of red blood cell. Depending on which globin or globins are underproduced, thalassemia can be further divided broadly into α, β, γ, δβ and εγδβ varieties. Among the different varieties, α- and β-thalassemia are the most common and clinically important types.

1.5 β-thalassemia

It is estimated that at least 80 to 90 million people, which constitute about 1.5% of the world's population are carriers of β-thalassemia. These carriers are mainly distributed in regions previously endemic for malaria, including the Mediterranean, Middle East, Africa, India, Southeast Asia and southern China. Although a large number of different β-thalassemia mutations have been identified, only a few common mutations and a varying number of rare ones account for most of the cases in those high frequency areas (Thein 1998).

β-thalassemia is caused by mutations that lead to the reduction or absence of β-globin chain. Generally, there are two main classes of β-thalassemia, namely the β^0 thalassemia, where totally no β-globin chain is being produced from the affected allele; as well as the β^+ or β^{++} thalassemia, in which a reduction of β-globin chain is being produced, in a severe or mild manner, respectively. A quantifiable deficiency of functional β-globin chains will lead to an imbalanced globin chain production and this results in an excess of α-globin chains (Weatherall and Clegg 2001; Schrier 2002). Excessive α-globin chains will aggregate in red cell precursors, forming inclusion bodies that cause an ineffective erythropoeisis (Hall and Thein 1994). Previous studies have shown that the severity of β-thalassemia is related to the degree of globin chain imbalance (Weatherall 1998).

Generally, β^0 thalassemias are caused by mutations that abrogated mRNA translation either at the initiation or extension phases of globin synthesis. These mutations normally affect the initiation codon or a splice junction site, resulting in a nonsense codon or frameshift mutation. On the other hand, β^+ thalassemias are caused by mutations affecting the transcription or mRNA processing (Thein 1998). Currently, there are at least fourteen deletions affecting the β-globin gene. Among these rare deletions, only a 619 bp deletion involving the 3' end of the β-globin gene is common in the Sind and Punjabi populations from India and Pakistan and this particular deletion accounts for 20% of the β-thalassemia alleles in these populations (Thein, Old et al. 1984; Varawalla, Old et al. 1991).

The clinical phenotypes of β-thalassemias are found to be vastly diversified. From symptomless to profound anemia which requires regular blood transfusion, they can be generally classified into three major groups, thalassemia major, thalassemia intermedia and thalassemia trait. Firstly, the most serious and fatal group, thalassemia major, also termed Cooley's anemia, describes a severe transfusion-dependent anemia. This type of thalassemia is normally present during the first year of life as the level of Hb F starts to decline from its initial levels after birth. When there is insufficient or inadequate blood transfusion, affected children will have symptoms of growth retardation, pallor, icterus and characteristic skeletal changes by progressive expansion of the bone marrow. This normally occurs when the affected child inherited two β-thalassemia genes in compound heterozygous or homozygous states.

On the other hand, β-thalassemia intermedia covers a wide range of clinical phenotypes from a condition slightly less severe compared to thalassemia major to a symptomless disorder that is only ascertained by routine examination of blood. It is found that the genotypes for β-thalassemia intermedia are very heterogenous, resulting from the interactions of one or two β-thalassemia alleles and other genetic variables (Thein 1998; Weatherall, Stamatoyannopoulous et al. 2001).

Lastly, β-thalassemia traits are characterized by a mild anemia with hypochromic microcytic red blood cells, low MCV (Mean Corpuscular Volume) and MCH (Mean Corpuscular Hemoglobin), increased level of Hb A_2 (3.5% to 5.5%) and slightly increased level of Hb F (less than 2%). Normally, they are in heterozygous state of β^0- or β^+- thalassemias.

To date, almost 200 β-thalassemia mutations have been characterized and reported. In contrast to α-thalassemias, β-thalassemias are rarely caused by major gene deletions and the majority of β-thalassemias are caused by point mutations within the β-globin genes. These point mutations result from single base substitutions, minor insertions or deletions of a few bases within the gene or its adjacent flanking sequences and they may affect any level of gene expression.

1.5.1 Dominantly inherited β-thalassemia

Typically, heterozygotes are clinically asymptomatic and the inheritance of two mutant alleles (as homozygotes) is required to produce a clinical disease. However, in some forms of β-thalassemia, the inheritance of a single β-thalassemia allele, in the presence of a normal complement of β-globin gene, may result in a clinically detectable phenotype (Thein 1999). This unusual form of thalassemia is termed the dominantly inherited β-thalassemia.

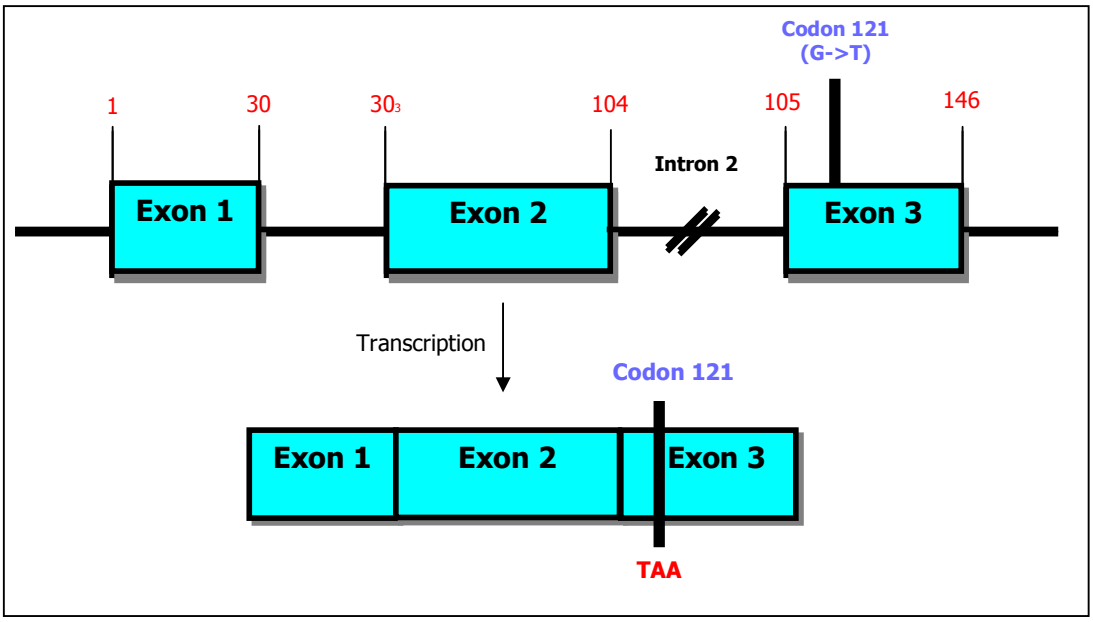
For heterozygous β-thalassemia, it normally features a thalassemia intermedia phenotype with moderate anemia, splenomegaly and a thalassemic blood picture. However, for the unusual dominantly inherited thalassemia, dyserythropoiesis associated with large intraerythroblastic inclusions are observed in addition to the usual features of heterozygous β-thalassemia (Thein 1998; Weatherall and Clegg 2001). Also, unlike the more common recessively inherited varieties, which lead to a quantitative reduction in normal β globin chain, they result in the synthesis of highly unstable β chain variants. It is even noted in many cases; these β chain variants are so unstable, that they are not detectable and only can be implicated from the DNA sequence.

More than 30 dominantly inherited β thalassemia alleles have been described (Thein, Hesketh et al. 1990; Thein 1999) and the mutations causing this unusual form of β-thalassemia include missense mutations, minor deletions, frameshifts resulting in elongated β variants and nonsense mutations resulting in truncated β chain variants. It has been reported that most of the nonsense mutations associated with dominant β-thalassemia are located in exon 3, whereas the majority of the nonsense mutations

associated with recessive β-thalassemia are usually located in exons 1 and 2 (Hall and Thein, 1994; Thein, 1999).

A very classic example of a mutation that leads to dominantly inherited β thalassemia is substitution of G to T mutation in codon 121 (Cd121). The G to T mutation is the very first nonsense codon mutation in exon 3 that has been reported to result in a premature termination of translation (Figure 4). It has been shown that minimal (0.05% to 0.1% of total) amounts of the truncated β globin chain were identified in a patient with codon 121 mutation (Adams, Steinberg et al. 1990) but in another study, it was shown that substantial amount of mutant mRNA (at least equal to normal β mRNA), are present in the reticulocytes of patients with Cd121 mutations, (Hall and Thein 1994). By putting these two findings together, it is enough to show that even with significant amount of β-121 mutant mRNA that is available for translation, the hyperinstability of truncated β-globin chains has resulted in the minimal amount of truncated β globin chains that are detectable. It is then postulated that the highly unstable β-globin chain variants will not be able to form functional tetramers and hence will precipitate intracellularly with the excess α-chains, forming inclusion bodies and finally lead to increased ineffective erythropoiesis (Thein, 1997).

Figure 4. Codon 121 mutation (premature termination codon at codon 121)



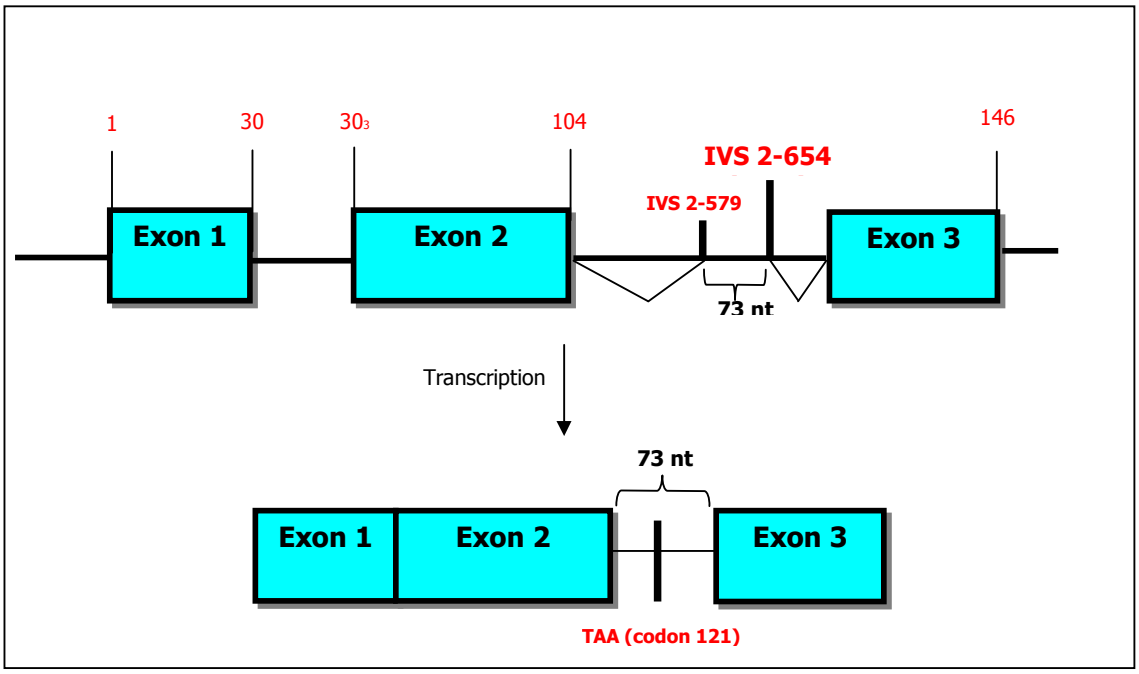
1.5.2 Cryptic splice-site mutations in introns

During mRNA splicing, the intervening sequences (IVS) (also known as introns), must be removed from the precursor mRNA, followed by the joining of coding regions to provide a functional template. In the splicing process, invariant dinucleotides of GT at the 5' (donor) and AG at the 3' (acceptor) splice junctions between exons and introns are the critical sequences. With the regions flanking both these invariant dinucleotides being well conserved, consensus sequences at exon and intron boundaries can be easily recognized. However, mutations occurring at these sites create a new cryptic splice site and hence leading to a reduction or complete inactivation of normal splicing.

To date, three β-thalassemia alleles having substitutions within intervening sequence (IVS) 2 of β-globin gene that generate new donor sites have been reported (Orkin et al. 1982; Dobkin et al. 1983; Cheng et al. 1984). Among these three

mutations, an interesting C to T mutation is present in the IVS 2, position 654 that is termed the IVS 2-654 mutation. This mutation is commonly found in the Asia population as a cause of recessive β-thalassemia (Chan, Chan et al. 1987). As a result of C to T substitution, this mutation creates a cryptic 5' donor splice site, which is spliced to the normal 3' acceptor site, while a cryptic 3' acceptor site is activated upstream at IVS 2-579 and spliced to the normal 5' donor site at the exon 2/intron 2 junction. Due to this cryptic splicing, an extra 73 nucleotides of intron 2 is incorporated into the aberrantly spliced mRNA transcript (Figure 5) resulting in a frameshift and premature termination at codon 121 (Ho, Hall et al. 1998).

Figure 5. IVS 2-654 mutation (premature termination codon at codon 121)



Although most of the heterozygotes for IVS 2-654 mutation give rise to a recessive phenotype, several cases with an unusually severe phenotype have also been reported (Naritomi, Nakashima et al. 1990). A father and a son, despite inheriting only a single copy of mutant allele, were found to manifest the phenotype of thalassemia intermedia. The ratio of aberrant to normal β-globin transcript was found to be 10 fold as compared to the asymptomatic heterozygotes for the same IVS 2-654 mutation (Ho, Hall et al. 1998). It was suggested that the severe phenotype might be caused by the accumulation of the aberrantly spliced mRNA, which presumably translates into a highly unstable β-globin chain variant. Interestingly, no abnormal protein was detectable in vivo in the study. These results implied that the more severe phenotype of IVS 2-654 maybe due to a second defect, possibly unlinked to the β-globin cluster that acts at the translational or post-translational level. (Ho, Hall et al. 1998)

1.6 Nonsense mediated decay (NMD)

Eukaryotic gene expression comprises a multistep pathway that is inter-connected; in which mRNA is the critical component. These steps include transcription, 5' cap formation, mRNA splicing, polyadenylation, mRNA export from nuclear pore to cytoplasm, mRNA translation and ultimately, mRNA degradation (Orphanides and Reinberg 2002). In order to ensure an efficient gene expression, cellular RNA are generally subjected to quality control or surveillance pathways that target premature termination (nonsense) codons (PTC) (Dimaano and Ullman 2004). One of the best characterized surveillance mechanisms is the nonsense mediated mRNA decay (NMD).

NMD is a post-transcriptional mechanism to control the quality of mRNA function by selectively degrading mRNAs that prematurely terminate translation due to the harboring of PTCs (Maquat and Carmichael 2001) (Maquat 2002). This quality control mechanism prevents the production of C-terminally truncated proteins that may cause deleterious dominant negative or gain of function effects (Frischmeyer and Dietz 1999). It is estimated that one third of inherited genetic diseases and many forms of cancers are due to PTCs, of which most of them are target of NMD (Holbrook, Neu-Yilik et al. 2004). This fact has pointed up the importance of NMD as the cell surveillance mechanism. In addition to the PTC-containing mRNA, NMD also targets on other abnormal transcripts which are a result of routine errors, such as alternative spliced mRNAs (Maquat 2002).

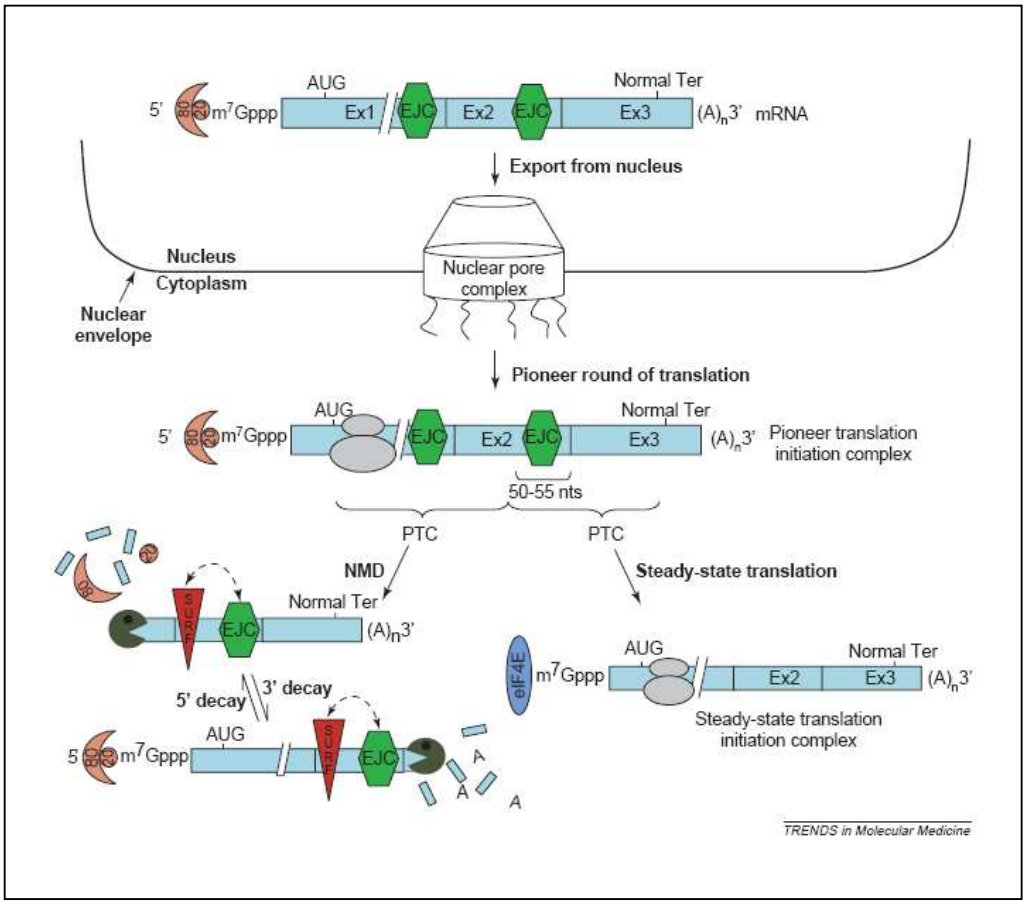
To date, the puzzle of how NMD recognizes PTC in mRNA transcript remains to be explored. However, in mammalian cells, even before the identification of the NMD molecular players, a general rule for the identification of PTCs that manage to trigger NMD was postulated; if PTCs are located more than 50-55 nucleotides upstream of the 3' most exon-exon junction, the mRNA will be subjected to NMD, whereas PTCs located downstream of this region will not be targeted (Li and Wilkinson 1998; Nagy and Maquat 1998; Thermann, Neu-Yilik et al. 1998). From this finding, it is clear that PTC recognition is dependent on the definition of exon-exon junctions, suggesting the role of mRNA splicing in NMD.

Besides that, NMD targets are recognized depending on a post-splicing exon junction complex (EJC) that is found deposited 20-24 nucleotides upstream of exon-exon junction (Le Hir, Izaurralde et al. 2000; Le Hir, Gatfield et al. 2001). During the

pre-mRNA splicing, EJC will be formed and functions to direct mRNA nuclear export and recruit up-frameshift (UPF) proteins (UPF-1, UPF-2, UPF-3 or UPF-3X) that are required for NMD (Lykke-Andersen 2001; Schell, Kulozik et al. 2002; Maquat 2004). EJC also serves as the binding platform for NMD factors (Le Hir, Gatfield et al. 2001).

In addition to that, later discoveries have found that the mammalian NMD is also triggered during the pioneer round of translation, which is the first time that mRNA passes through ribosome, with the mRNA is still bound to the nuclear cap-binding complex, CBP20 and CBP80 (Maquat 2004; Chang, Imam et al. 2007). After EJC is formed, newly spliced mRNAs are exported to the cytoplasm. Most of the mRNAs undergo pioneer round of translation during the process of export. Regardless of the location of pioneer round of translation occurs, if a PTC resides more than 50-55 nucleotides upstream of the last exon-exon junction (more than 20-25 nucleotides upstream of an EJC), the UPF-1 component of SURF complex (SMG-1, UPF-1, eRF-1 and eRF-30) will interact with UPF-2 on EJC and trigger NMD (Hosoda, Kim et al. 2005). NMD will then trigger the mRNA decay from one or both side of 5' and 3' ends (Lejeune, Li et al. 2003). If PTCs are found to be situated less than 50-55 nucleotides from the final exon-exon junction, NMD will not be triggered. The mRNA transcript will be channeled to the steady state translation initiation complex. The difference of steady state translation initiation complex with pioneer translation initiation complex is the presence of eukaryotic translation factor (elF-4E) rather than CBP-80 and CBP-20 complex at 5' cap region (Figure 6). When the first round of translation occurs, ribosome removes EJC or other mRNA ribonucleoprotein (mRNP) complexes from along the entire coding region till the termination codon (Baker and Parker 2004).

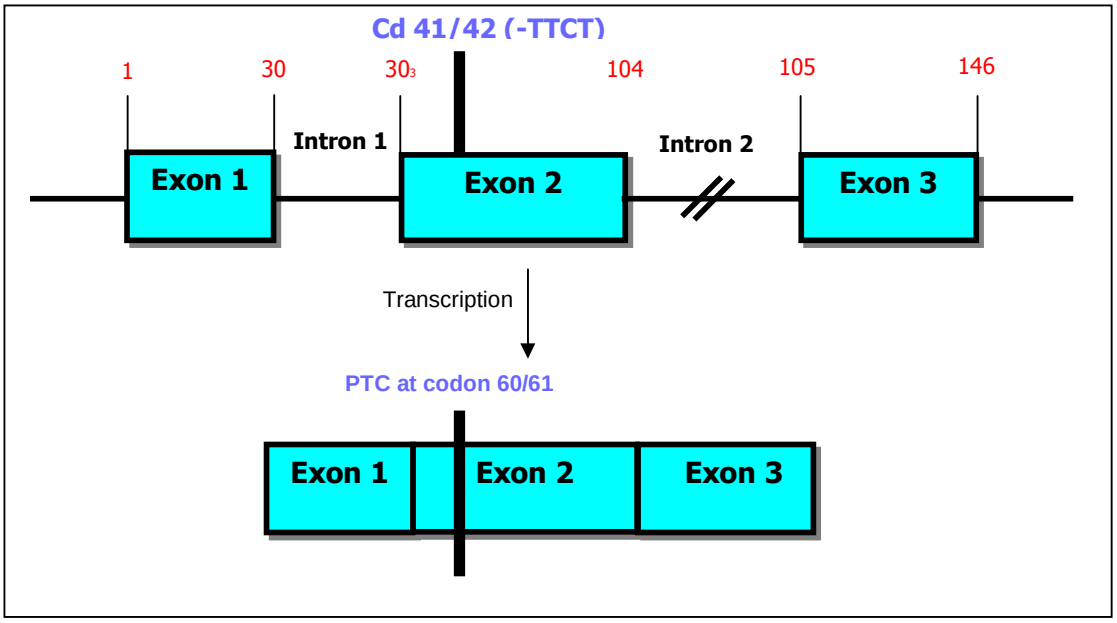
Figure 6. The NMD mechanism



1.7 NMD in β-thalassemia

In humans, NMD was initially discovered in the studies of β^0-thalassemia, which is caused by PTCs (Maquat, Kinniburgh et al. 1981; Kinniburgh, Maquat et al. 1982). Since then, the role of NMD as a modifier of the phenotype generated by nonsense mutations has become more evident (Frischmeyer and Dietz 1999). Among the β-thalassemia mutations, the majority generates premature termination codons in the first or second of the three exons in β-globin gene, producing a recessive phenotype, for example codon 41/42 mutation (Figure 7). Hence, individuals who are heterozygous for these mutations are generally asymptomatic and exhibit either absent or low levels of mutant β-globin mRNA (Baserga and Benz 1988). However, the nonsense mutations that located in exon 3 (the last exon) of the β-globin gene are associated with dominantly inherited β-thalassemia and exhibit high levels of mutant β-globin mRNA (Figure 8) (Thein, Hesketh et al. 1990). Translation of the mutant transcript produces truncated β-chains, which causes the dominantly inherited β-thalassemia via a dominant negative molecular mechanism.

Figure 7. Codon 41/42 mutation



Human β-globin transcripts with nonsense mutations in the last exon are resistant to NMD, a finding that is consistent with current knowledge of the mechanism of NMD. In addition, most human β-globin transcripts containing PTCs more than 50 nucleotides upstream of last exon-exon junction are subjected to NMD (Figure 9). Finer mapping shows that a boundary exists in exon 2 (Baserga and Benz 1988). However, it has been shown in erythroid cells that nonsense mutations in the 5' half of exon 1 are resistant to NMD, an exception for the current model of NMD (Romao, Inacio et al. 2000). Apart from the mutations downstream of the 3' boundary that are known to evade NMD, a few mutations in the 5' half of exon 1 have also been shown to be resistant to NMD. These observations suggest that the relative position of PTC to the last exon-exon junction is not sufficient to induce NMD in these cells.

Figure 8. Location of PTCs that contribute to the dominant and recessive phenotypes

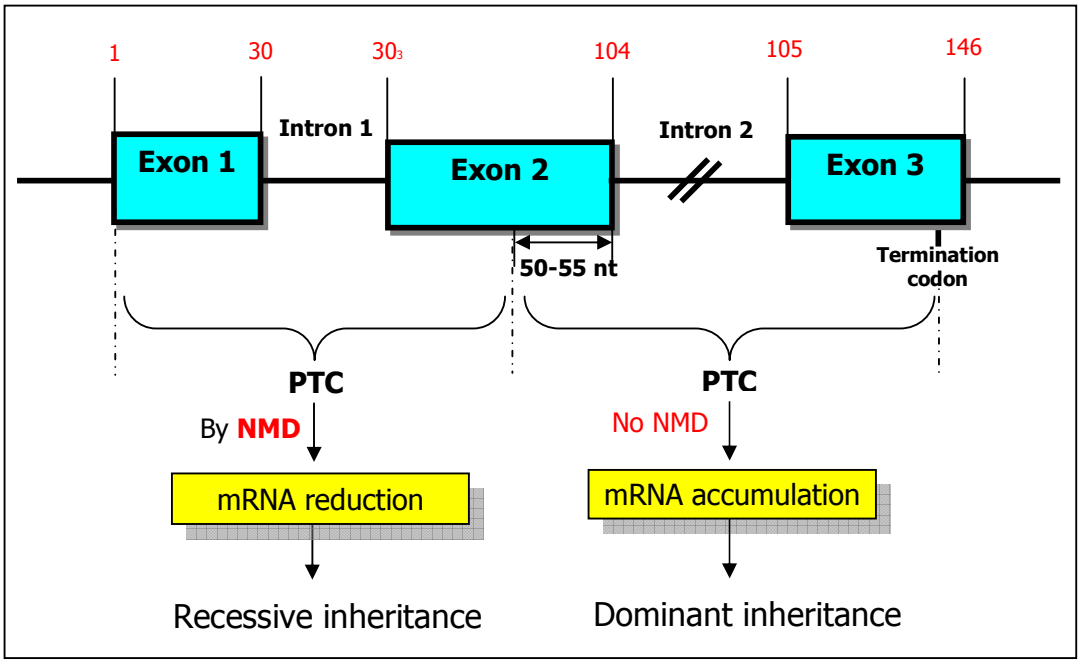
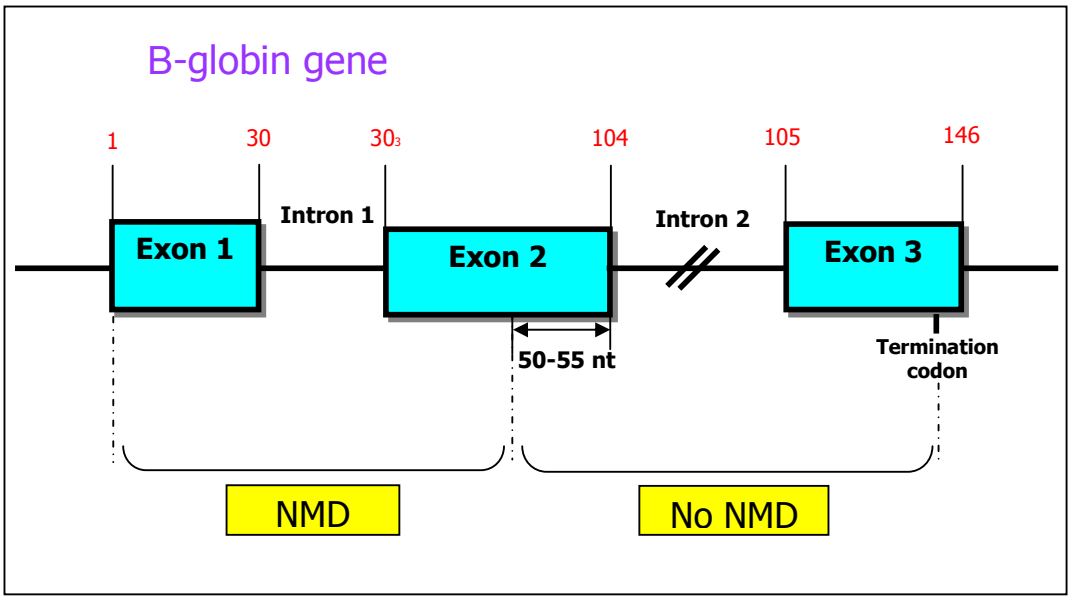


Figure 9. 50-55 nucleotide boundary rule of NMD



1.8 Aim of study

Since NMD was discovered, it has remained as a big puzzle, waiting to be solved. Though more and more studies have been reported on NMD, the actual mechanism of NMD in recognizing PTCs is still not clear. In this study, human β-globin was used as the platform to study the NMD mechanism. The main objective of this study is to elucidate the relationship between NMD with the location of PTCs in β-thalassemia. Though a general NMD rule, termed the 50-55 nucleotide boundary rule has been postulated, there are still some exceptions that do not apply to the rule. Hence, gene expression level of three mutations of β-thalassemia, codon 41/42 (-TTCT), codon 121 (G→T) and IVS 2-654 (C→T) that located in exon 2 and exon 3 were analyzed. We hypothesized that IVS 2-654 that contains extra 73 nucleotide intronic sequence is another exception of the 50-55 nucleotide boundary rule.

2.0 Materials and Methods

2.1 Plasmid construction

In this study, two main constructs were used for gene expression studies, namely the pHBB-EGFP and pTRE-HBB (Tet-Off) plasmids. Generally, both of these constructs contain the gene of interest, a 3.4 kb human β-globin gene fragment, comprising of a 1.6 kb β-globin gene, along with a 0.8 kb region of the beta globin promoter and the 5' UTR region, as well as a 1 kb fragment of the 3' UTR region which includes a downstream enhancer. All plasmids were constructed using the pBluescript (pBS) vector as the parental plasmid. As the main aim of this project is to study the effect of NMD on three different beta thalassemia mutations in comparison to the wild type human β-globin gene, mutant plasmids with the variant β-globin genes were constructed from the wild type β-globin gene by utilizing the in vitro mutagenesis method.

2.2 Generation of pBS-HBB mutant constructs

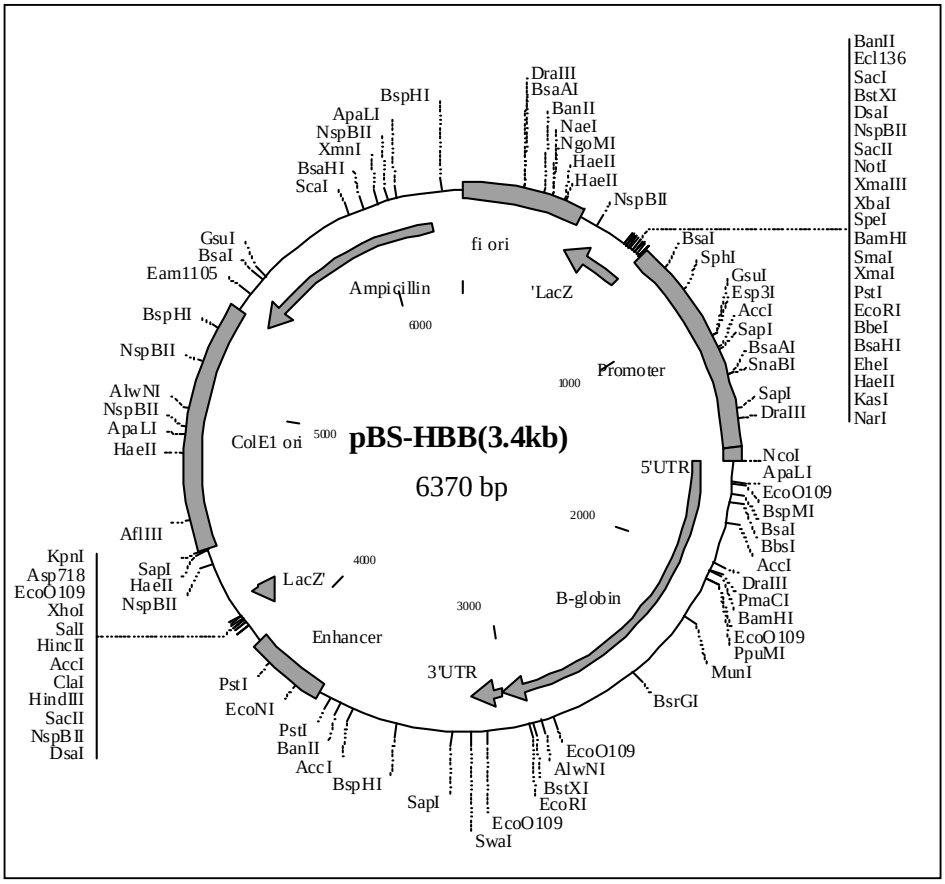
Three mutant constructs in either the pHBB-EGFP or pTRE-HBB plasmids were used in this study, namely codon 41/42, codon 121 and IVS 2-654 mutation constructs. Codon 41/42 mutation contains a deletion of 4 nucleotides (-TTCT) (Figure 7) whereas codon 121 mutation consists of a change of one nucleotide from G to T (Figure 4) and IVS 2-654 with a change of G to C (Figure 5). Construction of the codon 41/42 plasmids was performed by amplification of the β-globin gene containing this mutation from a patient's sample. For both codon 121 and IVS 2-654 plasmids, QuickChange II XL site directed mutagenesis kit was used to create the mutant constructs.

2.2.1 QuickChange II XL Site-Directed Mutagenesis

QuickChange II XL site-directed mutagenesis kit (Stratagene, cat no. 200522) is an in vitro mutagenesis kit that allows site-specific mutation in double-stranded plasmids to generate constructs with desired mutations for gene expression studies. In this study, different mutant constructs were created using plasmids containing the wild type human β-globin gene as the template plasmid.

For the mutant strand synthesis, two synthetic oligonucleotides containing the desired mutation, each complementary to opposite strands of the vector were used to generate mutant plasmids using the supercoiled double-stranded DNA vector (pBS-HBB) (Figure 10) as a template. These specific oligonucleotides were extended during temperature cycling by *PfuUltra*^{TM} HF DNA polymerase, without primer displacement. The reaction was carried out in a final volume of 50 µL, which contained 5 µL of 10X reaction buffer, 10 ng of dsDNA template (pBS-HBB), 125 ng of each mutagenesis primers, 1 µL of dNTP mix, 3 µL of QuickSolution, 1 µL of PfuUltra HF DNA polymerase (2.5 U/µL) and double distilled water. The thermal cycling reaction was performed in a Biometra T3 thermal cycler (Biometra, GmbH, Germany), with initial enzyme activation at 95 ºC for 1 minute, followed by 18 cycles of 95 ºC for 50 seconds, 60 ºC for 50 seconds and 68 ºC for 6 minutes. Subsequently, the reaction was incubated at 68ºC for 7 minutes. Following the temperature cycling, the reaction tube was placed on ice for 2 minutes to cool to ≤ 37ºC.

Figure 10. pBluescript-HBB plasmid (pBS-HBB)



Simultaneously, a control reaction was also included for each sample reaction. The content of the control reaction was similar with the sample reaction, with the exception that an optimized reaction with specific template vector and primers were used as a positive control for the mutagenesis reaction. For the control reaction, a 10 ng of pWhitescript 4.5-kb plasmid (5ng/µL) was used as the template vector and for the control primers, a pair of 34-mer primers provided with the kit, were used. In addition to that, the cycling parameters were slightly different compared to the sample reaction. The extension time for the control reaction was set at 68ºC for 5 minutes and the cycling was carried out for only 12 cycles.

Next, to digest away the parental supercoiled DNA vector and leaving only the newly synthesized vectors with mutations of interest, both sample and control reactions were digested with the *Dpn* I enzyme. *Dpn* I specifically recognizes and digests methylated and hemimethylated DNA. Therefore, to remove methylated template vectors, 1 µL of the *Dpn* I restriction enzyme (10U/µL) was added directly into the reaction tube and mixed gently by pipetting the solution up and down for several times. Following that, the reaction mixtures were then spun down for 1 minute and then incubated immediately at 37 ºC for 1 hour.

After the *Dpn* I digestion step, the mutant construct was transformed into the XL10-Gold Ultracompetent Cells which were thawed gently on ice before transformation. For each control and sample reaction, a total of 45 µL of the ultracompetent cells was aliquoted to a pre-chilled 15 mL falcon round bottom tube. After that, 2 µL of β-ME (β-mercaptoethanol) mix provided with the kit was added to the ultracompetent cells. The transformation reaction was then swirled gently to mix and incubated on ice for 30 minutes. Subsequently, the reaction tube was heat-pulsed at 42 ºC for 30 seconds and incubated on ice again for another 2 minutes. After the heat-shock step, a total of 0.5 mL of preheated (42 ºC) NZY^{+} broth was added to each tube. The reaction tubes were then incubated at 37 ºC for 1 hour with shaking at 225 - 250 rpm.

Following that, the samples and control transformation reactions were then plated on LB-ampicillin agar plates and incubated at 37 ºC for at least 16 hours. A total of 5 single colonies were picked from each plate using sterile small micropipetter tips and transferred into different 0.2 mL tubes, which contain 10 µL of LB medium

(1% Tryptone, 0.5% Yeast extract, 1% NaCl, pH 7.5). After that, these LB media that contain bacteria were used to inoculate 5 mL of LB medium with ampicillin (100 µg/mL) in a 15 mL Falcon tube. The tubes were shaken at 225 rpm at 37 ºC for overnight.

Plasmids (pBS-HBB) that contain HBB mutations were then isolated from bacteria culture by using the Wizard Plus Minipreps DNA purification system (Promega, cat No. A7510). The concentrations of plamids were determined by a spectrophotometer. After obtaining the required mutant plasmids, sequencing reactions were carried out in order to confirm the correct mutant sequences in the pBS-HBB plasmid. The sequencing reaction was carried out in a total volume of 20 µL reaction, containing 4 µL of BigDye Terminator Ready Reaction mix (Applied Biosystems, cat No. 4390242), 3.2 pmol of one primer and a total of 300 ng of plasmid DNA. A combination of 8 primers were used in the different sequencing reactions in order to make sure the full sequence of human β-globin gene can be checked. (Table 1)

Table 1. Primers sequences for sequencing HBB fragment

Name	Sequence 5' to 3'	Nucleotides ID in HUMHBB	Region	Length
61782-R	GTG GAA GAG CTT TGT CTA CC	61782-61763	5' UTR	20
62010-F	ACG GCT GTC ATC ACT TAG AC	62010-62029	5' UTR	20
Cd8/9-R	GGG CAG TAA CGG CAG AC	62229-62213	Exon 1	17
Cd41/42-F	CCT TGG ACC CAG AGG TT	62425-62441	Exon 2	17
InII654-F	TGA TAA TTT CTG GGT TAA GG	63265-63284	Intron 2	20
63466-R	AAG AGG TAT GAA CAT GAT TAG C	63466-63445	Intron 2	22
64093-F	TCT CTT GCT TAG AGA TAC CAC	64093-64113	3' UTR	21
64299-R	CAG ATT CCG GGT CAC TGT G	64299-64281	3' UTR	19

The sequencing reactions were performed in a Biometra T3 thermal cycler (Biometra, GmbH Germany), with 25 cycles of incubation at 96 ºC for 30 seconds, 50 ºC for 15 seconds and 60 ºC for 4 minutes. The extension products of sequencing were then purified by using ethanol/sodium acetate precipitation method. The purified DNA pellets were re-suspended in a total of 12 µL Hi-Di^{TM} Formamide (Applied Biosystems, cat No. 4311320) and transferred onto a 96-well reaction plate. The samples were denatured at 95 ºC for 2 minutes and analyzed by automated capillary electrophoresis on an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, USA). The rapid sequencing protocol, which only requires a 40 minutes electrophoresis through Performance-Optimized-Polymer (POP)-4^{TM} (Applied Biosystems), across a 36 cm long capillary, was used.

The sequences of mutant β-globin gene in pBS-HBB clone were aligned against the GeneBank HBBHUM sequence using Vector NTI Suite 7 Software. Clones with the correct sequence were selected and named as pBS-HBB codon 41/42, pBS-HBB codon 121 and pBS-HBB IVS 2-654.

2.3 Generation of pHBB-EGFP constructs

As previously mentioned, pHBB-EGFP construct is one of the main plasmids that were used in this study. To construct this plasmid, pBS-HBB with wild type and different mutations were used as the template plasmid, along with a pMDR-EGFP construct (MDR stands for multi drug resistance gene) (kindly provided by A/P Caroline Lee) (Figure 11). In short, the human β-globin gene from pBS-HBB construct was digested out and then ligated with the EGFP (enhanced green

fluorescent protein) fragment, using pMDR-EGFP as the backbone of the pHBB-EGFP construct. The EGFP gene in the pHBB-EGFP construct will serve as the reporter gene to determine transfection efficiency.

From the pBS-HBB constructs (wild type or with mutations) that were generated, the fragment of 3.4-kb β-globin gene was digested out, using restriction enzyme *Nar* I and *Sac* II, in a two separate digestion reactions. The first part of digestion was done in a total volume of 50 µL reaction, with 3 µg of each pBS-HBB plasmid DNA (wild type or with mutations) and 20 units of *Nar* I enzyme (4U/µL, New England Biolabs) in 1 x NE Buffer. The reaction was incubated at 37 ºC for 5 hours. Each linearized product was then gel purified and eluted with 30 µL TE buffer. Following by that, the linearized pBS-HBB fragment was digested for second time, using 10 units of *Sac* II enzyme (10U/µL, MBI Fermentas), with 1 x Buffer B^+ in a 50 µL reaction. After the incubation at 37 ºC for 5 hours, the final digestion product produced 3 fragments with different length: 3397 bp, 2916 bp and 57 bp, after the digestion product was separated by gel electrophoresis. The 3397 fragment that contained the human β-globin gene was cut out from the gel and purified using GFX^{TM} PCR DNA and gel band purification kit (Amersham Biosciences, cat No. 27-9602-01).

On the other hand, the digestion of pMDR-EGFP was also performed in two different reactions. Firstly, 5 µg of pMDR-EGFP were digested with 50 units of *Sac* II restriction enzyme (10U/µL, MBI Fermentas) in 1 x Buffer B^+, of a total 50 µL reaction volume. The digestion reaction was incubated for 3 hours at 37 ºC, generating 4 fragments with different sizes: 5990 bp, 3480 bp, 2780 bp and 30 bp.

The digestion product was separated using gel electrophoresis and the 5990 bp fragment that contained the MDR promoter, EGFP gene and Ampicillin resistance gene was gel purified and eluted in 30 µL of TE buffer. The 5990 bp fragment was subsequently further digested by using 20 units of *Nar* I restriction enzyme (4U/µL, New England Biolabs), with the addition of 1 x NE Buffer 1, in a 50 µL reaction. The digestion reaction was carried with the incubation for 3 hours at 37 ºC and the digested product was separated by gel electrophoresis. After the second digestion, two fragments were generated: 4448 bp and 1542 bp. The 4448 bp fragement containing EGFP gene and Ampicillin resistance gene was gel purified as described previously.

From the digestions from pBS-HBB (wild type or with mutation) and pMDR-EGFP, purified fragments of human β-globin gene (3397 bp) and EGFP (4448 bp) were obtained. These two fragments were then ligated to form the pHBB-EGFP construct (Figure 11). Subsequently, 20 ng of both HBB and EGFP fragments were mixed in the ligation reaction, with the addition of 2.5 units of T4 Ligase (5U/µL, MBI Fermentas) and 1 x ligation buffer. The ligation reaction was incubated at 22 ºC for 4 hours, followed by deactivation of T4 ligase at 65 ºC for 10 minutes.

Figure 11. Cloning strategy for pHBB-EGFP

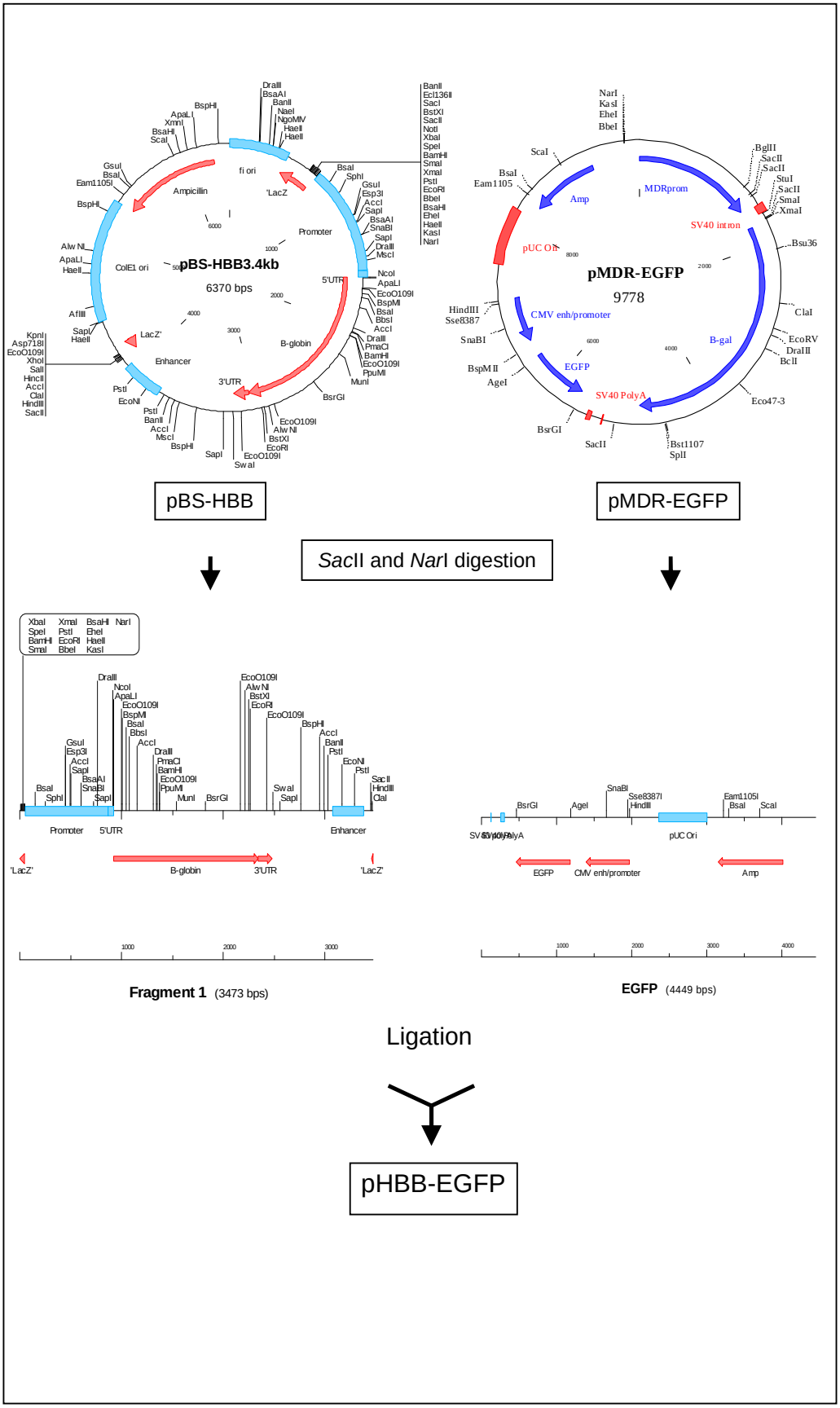
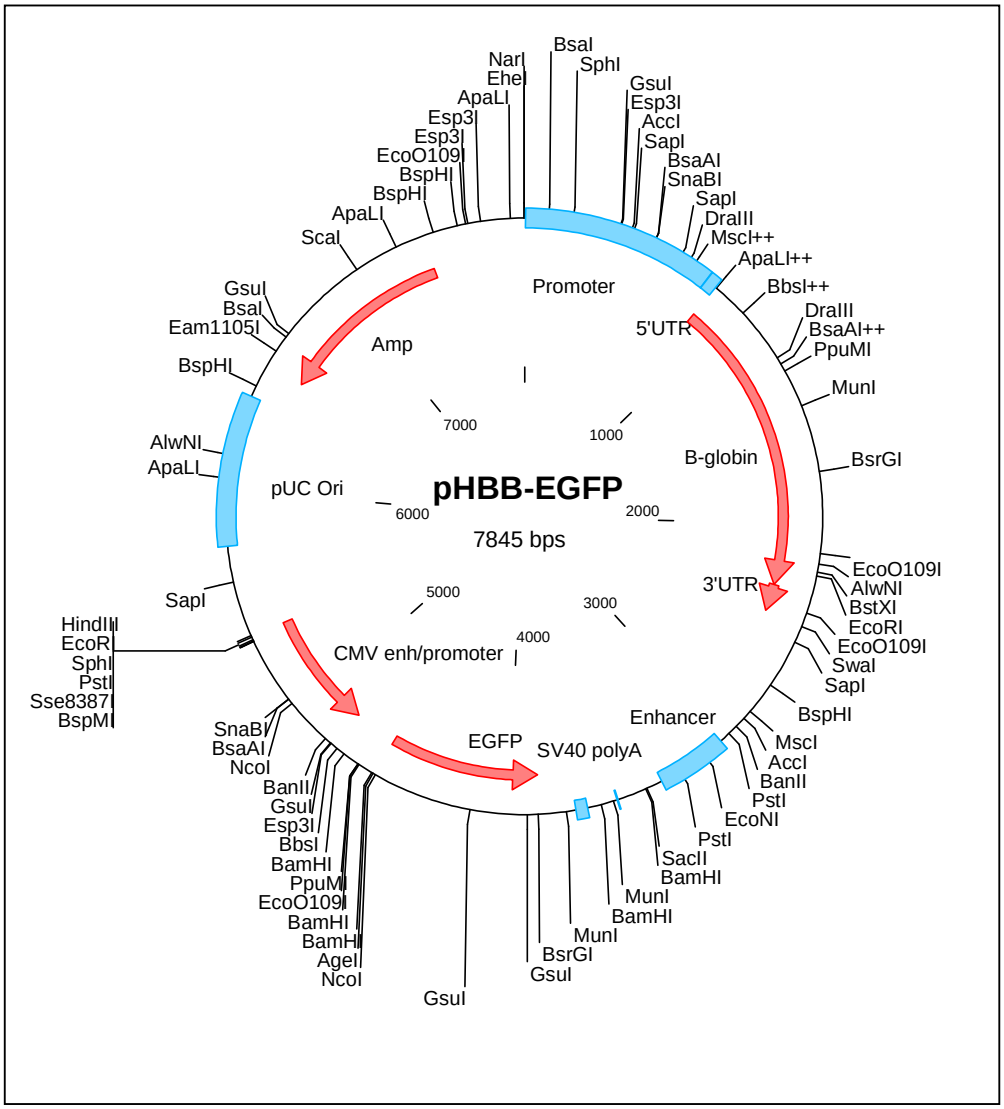


Figure 12. pHBB-EGFP plasmid



The pHBB-EGFP plasmids (wild type or with mutation) (Figure 12) were then transformed into the Subcloning efficiency™ DH5α™ competent cells (GibcoBRL Life Techonologies, cat No. 182655-017), using the protocol that was described previously. Five colonies from each plasmid were picked and used to inoculate in a 5 mL LB medium with ampicillin (100 µg/mL). The medium with bacteria were then shaked for 14 hours at 37 ºC. After the shaking, pHBB-EGFP (wild type or with mutation) was purified from the bacteria culture by using Wizard Plus Minipreps

DNA Purification System (Promega, cat No. A7510), using the manufacturer's protocol. Subsequently, plasmids concentrations were determined using spectrophotometer.

All the pHBB-EGFP (wild type or with mutation) plasmids were then checked for the correct fragment size by using restriction enzyme digestion. Each digestion reaction was carried out in a total of 20 µL reaction volume, with the addition of 20 units of EcoRI enzyme, 1 x buffer EcoRI^{+} and 1 µg of each plasmid DNA. The digestion reaction was incubated for 3 hours at 37 ºC. Digested products were then separated by gel electrophoresis and the correct recombinant plasmids that were digested, resulted in two digested fragments: 4695 bp and 3150 bp. For each correct pHBB-EGFP (wild type or with mutation) construct, maxiprep purification was carried out by using QIAGEN EndoFree^{TM} Plasmid Maxi purification kit, to prepare for the transfection purpose.

2.3.1 HeLa steady state transient transfection

HeLa cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (Gibco), supplemented with 10% Fetal Calf Serum (FCS). Before seeding cells, HeLa cells were checked under microscope to ensure the cells are in good condition for plating. After that, they were trypsinized with 1 mL of 1 x Trypsin and resuspended in 8 mL of fresh medium. Cells were counted using hemocytometer and seeded in a 6-well plate at a density of 5 x 10^5 cells per well. For each well, the volume was topped up with 3 mL of DMEM media without FCS. Cells were then incubated for 24 hours at 37 ºC.

After 24 hours of incubation, the cells were checked under microscope to make sure that the cells have reached a confluency of about 90% suitable for transfection. Following that, the Lipofectamine 2000 reagent was prepared by diluting 10 µL in 250 µL of DMEM without FCS for each well. The reaction mixtures were then left at room temperature for 5 minutes. During the incubation time, the plasmid DNAs were prepared by diluting 4 µg of pHBB-EGFP (wild type or with mutation) in 250 µL of DMEM without FCS for each well. The diluted DNAs were then combined with the diluted Lipofectamine 2000 and mixed gently. The reaction mixtures were incubated for 20 minutes at room temperature. After the incubation step, the complexes were added into each well and the plates were mixed gently by rocking them back and forth. One well was left as negative control (without transfection). The plates were incubated for 24 hours at 37 ºC in the CO_2 incubator.

The following day, transfected cells were checked under fluorescent microscope with FITC filter for transfection efficiency. Fluorescence signals were observed in transfected cells and the transfection efficiency for each well was determined by comparing the number of transfected cells with fluorescence signal to the untransfected negative control cells without fluorescence. The transfection efficiency was determined to be at least 50-60% before proceeding to the harvesting step. Cells were then harvested by adding the TRIzol and the lysates were kept in −86ºC freezer prior to RNA extraction step.

2.4 Generation of pTRE-HBB constructs

In the BD^{TM} Tet-Off gene expression kit, the pTREhyg construct was provided as the response plasmid, which expresses the gene of interest under control of the tetracycline response element (TRE). In the construction of pTRE-HBB expression vector, the pTREhyg response vector was digested with restriction enzymes *Not* I and *Sal* I, to generate a fragment of 5306 bp. This fragment contained the TRE element and the hygromycin resistance gene. The digestion reaction was carried out in a 60 µL reaction with the addition of 1 x Buffer Y^{+}/TANGO^{TM} , and incubated at 37 ºC for 3 hours. The digested fragment was then separated by gel electrophoresis to check for the correct fragment. Subsequently, the desired fragment was gel purified and eluted in 30 µL of TE buffer (Figure 13).

On the other hand, pBS-HBB (Wild type or with mutation) was also digested with restriction enzymes *Not* I and *Sal* I, to generate a HBB fragment of 3473 bps. The digestion condition was exactly the same with what was described above for pTREhyg construct, as both were using the same restriction enzymes. Subsequently, the fragment with the size of 3473 bp was gel-purified and eluted in a total of 30 µL TE buffer.

From the digetions from pBS-HBB (wild type or with mutation) and pTREhyg, purified fragments of human β-globin gene (3473 bp) and TRE (5306 bp) were obtained. These two fragments were then ligated to form the pTRE-HBB construct (Figure 13). A total of 20 ng HBB and TRE fragments were mixed respectively in the ligation reaction, with the addition of 2.5 units of T4 Ligase (5U/µL, MBI Fermentas)

and 1 x ligation buffer. The ligation reaction was incubated at 22 ºC for 4 hours, followed by deactivation of T4 ligase at 65 ºC for 10 minutes.

Figure 13 Cloning strategy for pTRE2-HBB (Tet-Off)

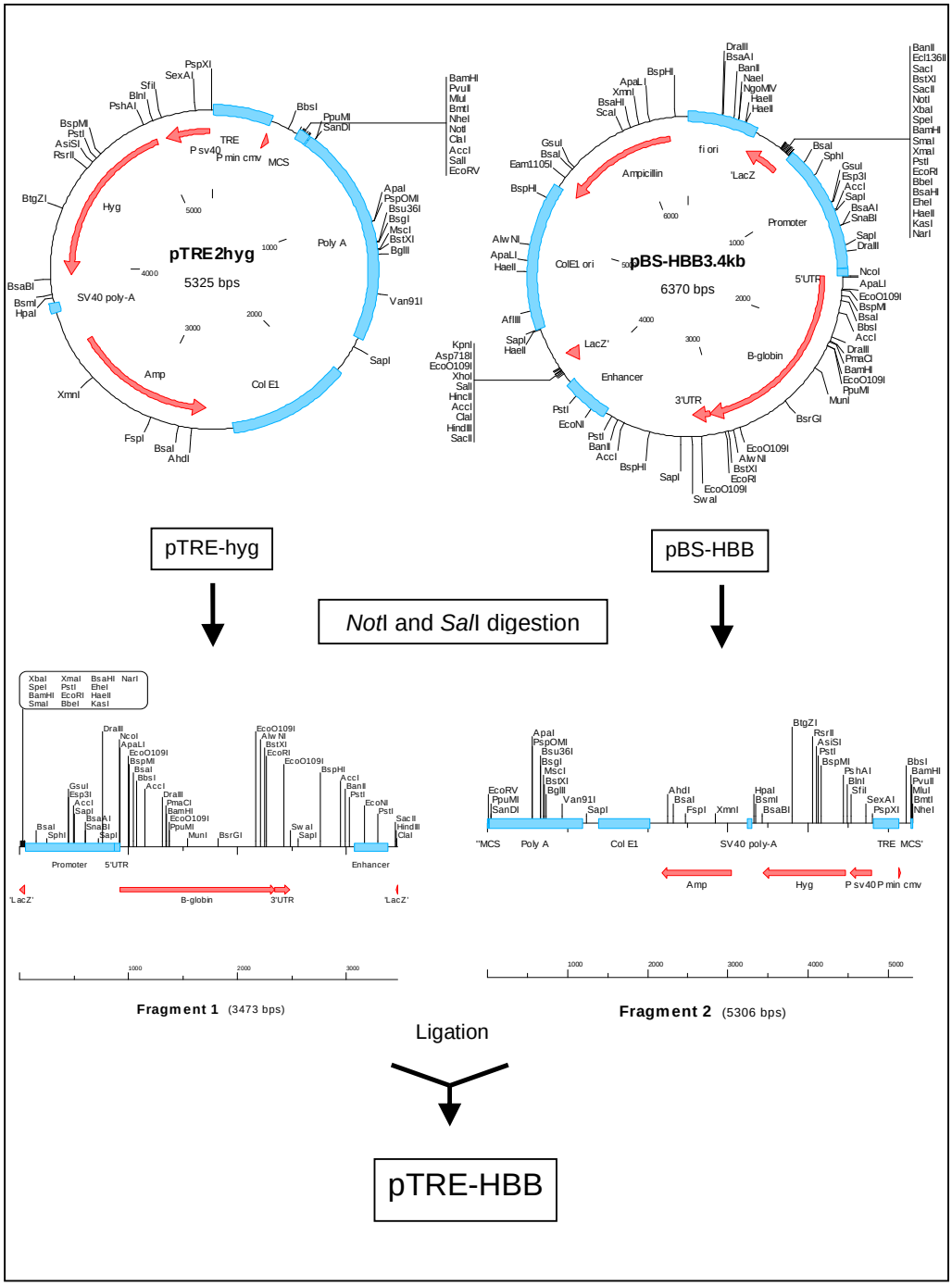
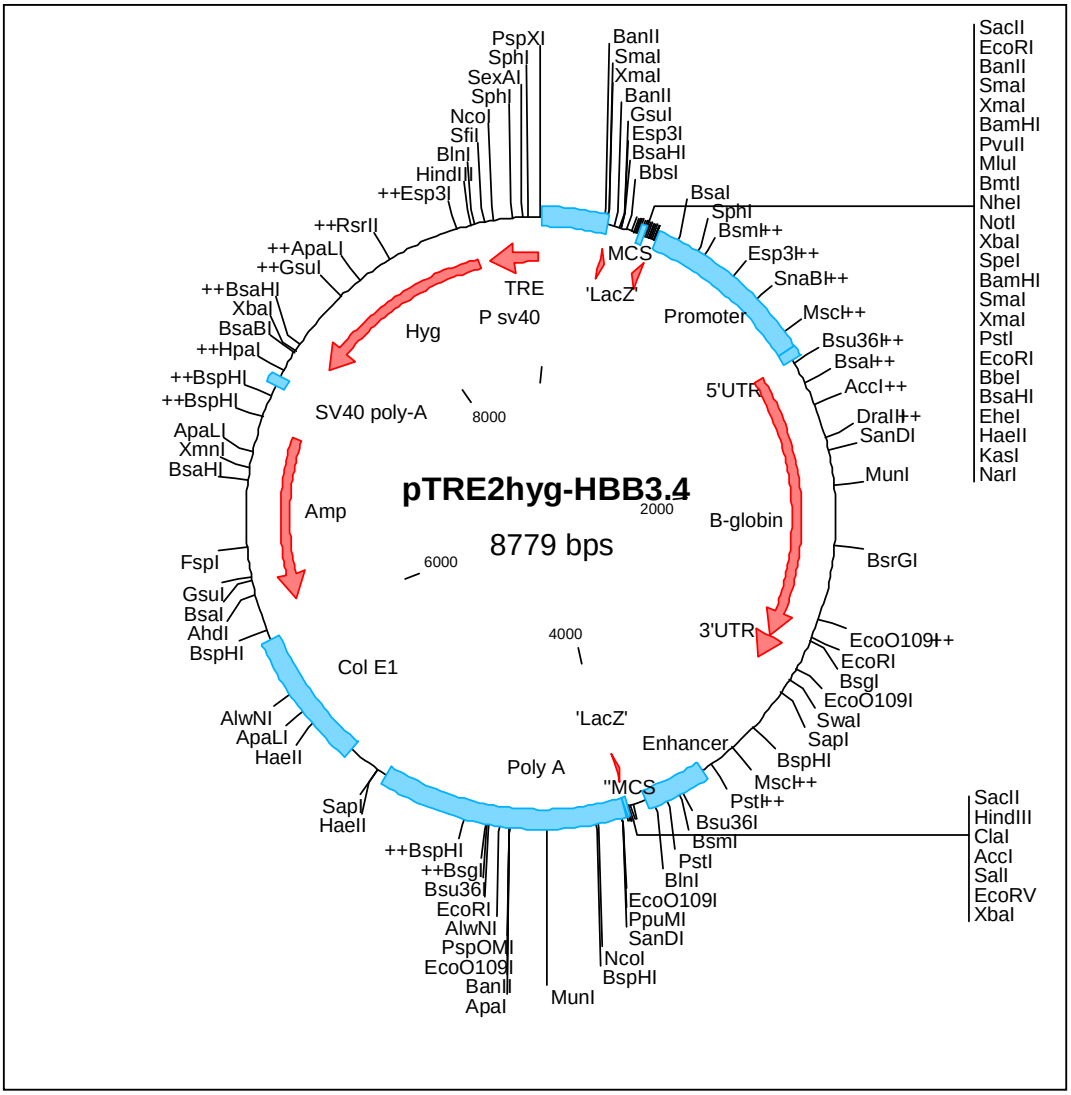


Figure 14. pTRE-HBB plasmid



pTRE-HBB (wild type or with mutation) (Figure 14) was transformed into the Subcloning efficiency™ DH5 α™ competent cells (GibcoBRL Life Techonologies, cat No. 182655-017) as described previously in the last section. After determining the correct plasmid, purification was carried out using QIAGEN EndoFree™ Plasmid Maxi purification kit, to prepare for the transfection purpose. Purified plasmids were then quantitated using spectrophotometer.

2.4.1 BD^{TM} Tet-Off system

In Tet-Off system, gene expressions are turned on when tetracycline or doxycycline is removed from the culture medium. Hence, this system enables gene expression to be tightly regulated in response to varying concentrations of tetracycline or doxycycline. The first critical component of Tet-Off system is the regulatory protein which contains Tet-repressor (TetR) and a VP16 activation domain. Addition of VP16 domain will convert the TetR from a transcriptional repressor to a transcriptional activator and the resulting hybrid protein is known as the tetracycline-controlled transactivator (tTA).

The second critical component is the response plasmid which functions to express the gene of interest (gene X) under control of the tetracycline-response element (TRE). When cells contain both regulatory (pTet-Off) and the response (pTRE-geneX) vectors, the gene X will be expressed upon the binding of tTA protein to the TRE and activates transcription in the absence of doxycycline.

2.4.2 mRNA decay study using HeLa Tet-Off system

As mentioned in the Tet-Off system, the two critical components are the regulatory protein (pTet-Off) and the response plasmid (pTRE-gene X). The gene of interest in this study is the human β-globin gene, hence the response plasmid constructed were named as pTRE-HBB (wild type or other mutation), as previously described. As for the regulatory component, a HeLa Tet-Off cell line was purchased from BD Biosciences Clontech in order to reduce the time needed for establishing a Tet-Off cell line. HeLa Tet-Off cells are similar to ordinary HeLa cells except that

these cells can stably express tTA from the integrated copies of pTet-Off (the regulator plasmid). HeLa Tet-Off cells were cultured using Dulbecco's modified Eagle's medium (DMEM) (Gibco), supplemented with 10% of Tet Fetal Calf Serum (Tet-FCS). This cell line was cultivated in an atmosphere of 5% CO_2 at 37 ºC and routinely passaged by trypsinization.

Before seeding cells, HeLa Tet-Off cells were checked under microscope to ensure the cells are in good condition for plating. Subsequently, they were trypsinized with 1 mL of 1 x Trypsin and resuspended in 8 mL of fresh medium. Cells were counted using hemocytometer and seeded in a 6-well plate at a density of 5 x 10^5 cells per well. A 5-time-point experiment was designed to study the mRNA decay rate, hence for each pTRE-HBB (wild type or with mutation), cells were seeded in 5 wells and supplemented with fresh medium with FCS. One well will serve as the non transfected negative control. Subsequently, the plates with cells were incubated at 37ºC for 24 hours.

Twenty-four hours after cell seeding, each well is expected to have reached 90% of cell confluency. The cells in the 6-well plates were checked again before proceeding with transfection. Subsequently, Lipofectamine 2000 was prepared by diluting 10 µL in 250 µL of DMEM without FCS for each well and incubated at room temperature for 5 minutes. During the incubation time, the plasmid DNA mix was prepared. For each well of transfection reaction, 3 µg of pTRE-HBB and 1.5 µg of pCMV-EGFP were prepared in 250 µL of DMEM without FCS. The pCMV-EGFP construct was co-transfected to serve as a control to ensure equal transfection efficiency and to show equal amplification efficiency in the real time assays later.

Following by that, the diluted DNAs and was combined with the diluted Lipofectamine 2000. The reaction mixtures were mixed gently and incubated for 20 minutes at room temperature. After the incubation, the mixtures were added into each well and the plates were mixed gently by rocking them back and forth. The plates were incubated for 24 hours at 37 ºC in the CO_2 incubator. Due to toxicity of Lipofectamine, after 4-5 hours of incubation, cells in all the wells were washed by 1 x PBS buffer and replaced with fresh medium, supplemented with 10% Tet-FCS.

After 24 hours of incubation, the transfection efficiency of the HeLa Tet-Off cells was checked under fluorescence microscope with FITC filter, by comparing the transfected and untransfected wells. Following by that, cells were harvested from the first well of 6-well plate, which was designated for 0 hour time point, without adding Doxycycline. Cells were harvested by adding 1 mL of TRIzol directly into the well and incubated for 5 to 10 minutes. During the incubation time, 100 ng/mL of Doxycycline was added into the remaining 5 wells on the plate and were incubated at 37 ºC in CO_2 incubator. For the harvested cells, the lysate with TRIzol was kept in a 2 mL tube and kept in -86 ºC freezer before the RNA extraction step.

The remaining wells of cells in the plate were then harvested according to the time point of 4 hours, 8 hours, 12 hours and 24 hours after the addition of Doxycyline. Similarly, all the cells were harvested by adding 1 mL of TRIzol and kept in -86 ºC freezer before the RNA extraction step.

2.5 Development of MEL stable cell lines

The development of MEL stable cell lines was carried out by using the pHBB-EGFP wild type construct or with mutations (codon 41/42, codon 121 and IVS 2-654) and pNeo (plasmid containing the neomycin resistance gene). Firstly, pHBB-EGFP constructs was digested by *Sca* I restriction enzyme in order to linearize the circular plasmid. The digestion reaction was carried out in a total volume of 100 µL reaction. 15 µg of maxiprep pHBB-EGFP plasmids were digested by 60 units of *Sca* I restriction enzyme (10 U/ µL, MBI Fermentas) in 1 X supplied Buffer K and also 0.1 µg/µL of acetylated BSA. HPLC water was added to top up the reaction volume to 100 µL. The digestion reaction was incubated at 37 ºC for 3 hours. Subsequently, 5 µL of digestion reaction was used to run on gel electrophoresis to check if the digestion was successful. The remaining 95 µL of digestion product was then purified by using GFX^{TM} PCR DNA and gel band purification kit (Amersham Biosciences, cat No. 27-9602-01) and eluted in TE buffer, with a final volume of 30 µL.

Subsequently, pNeo construct was digested by using *Alw44* I in order to linearize the circular plasmid. The digestion reaction was carried out in a total volume of 40 µL reaction. 9.5 µg of maxiprep pNeo plasmids were digested with 50 units of *Alw44* I restriction enzyme (10 U/ µL, MBI Fermentas) in 1 X Buffer Y^{+}/TANGO. HPLC water was added to top up the reaction volume to 40 µL. The digestion reaction was incubated at 37 ºC for 3 hours. Subsequently, 2 µL of digestion reaction was used to run on gel electrophoresis to check if the digestion was successful. The remaining 38 µL of digestion product was then purified by using

GFX^{TM} PCR DNA and gel band purification kit (Amersham Biosciences, cat No. 27-9602-01) and eluted in TE buffer, with a final volume of 30 µL.

After linearizing both the pHBB-EGFP (wild type or with mutation) and pNeo plasmids, MEL cells were then seeded in a 24-well plate, with the density of 3 x 10^5 cells per well in 500 µL of ATCC RPMi 1640 medium without neomycin. The MEL cells were then incubated for 24 hours before the transfection step.

In the transfection procedure, the linearized pHBB-EGFP and pNeo plasmids will be co-transfected into the MELcells using Lipofectamine 2000 (Invitrogen). The transfection mixture was prepared by diluting 1 µg of pHBB-EGFP and 0.03 µg of pNeo contructs in 50 µL of ATCC RPMi 1640 media without serum. The reaction was mixed gently and thoroughly. On the other hand, 2 µL of Lipofectamine 2000 was diluted in a 50 µL of ATCC RPMi 1640 and incubated for 5 minutes at room temperature. After the 5 minutes incubation, both the diluted plasmid DNAs and Lipofectamine 2000 were combined and made up to a total of 100 µL reaction. The reaction was mixed gently and incubated at room temperature for 20 minutes.

After 20 minutes of incubation, the transfection mixture which contains plasmid DNAs and transfection reagent (Lipofectamine 2000) was then added into each well containing MEL cells and ATCC RPMi 1640 medium. With the incubation for 24 hours, the MEL cells density was expected to reach about 90-95% confluency. After adding the transfection mixture, the plate was mixed gently by rocking back and forth for a few times.

Subsequently, the 24-well plate was incubated at 37 ºC in a CO_2 incubator for 24 hours. The transfected cells were checked under fluorescence microscope for the transfection efficiency. Following by that, the cells were passaged at the ratio of 1:10-12 into fresh growth medium, without neomycin and incubated at 37 ºC for another 24 hours. The subsequent day, which is 48 hours after transfection, 1 mg/mL of neomycin (G418) was added into the cells. For every 3 to 4 days, medium was changed and replenished with fresh medium with 1 mg/mL of G418.

About 15 days later, with the frequent passaging and changing of fresh medium with G418, the selection was complete and isolated colonies began to appear. The surviving cells were those with fluorescence and able to survive with the presence of G418. The suspension culture was then diluted and large, healthy colonies were chosen and transferred to individual wells. From this point onwards, the concentration of G418 was reduced and maintained at 400 µg/mL. These stable cell lines were then cultured in ATCC RPMi 1640, with 10% of un-inactivated fetal calf serum (FCS) and 400 µg/mL of G418.

2.5.1 Erythroid differentiation induction of MEL stable cell lines

MEL stable cell lines were maintained in ATCC RPMI 1640 medium supplemented with 10% of Fetal Calf Serum (FCS) and 400 µg/mL of G418. All the cell lines were cultivated in an atmosphere of 5% of CO_2 at the temperature of 37 ºC.

In this erythroid differentiation induction experiment, erythroid cell differentiation was induced in an equal amount of MEL stable cell lines by adding 2% (vol/vol) of dimethy sulfoxide (DMSO) to the media. Firstly, MEL stable lines

(Norm, codon 41/42, codon 121 and IVS 2-654) were checked under fluorescent microscope with FITC filter to ensure all the cells were fluorescing and in healthy condition. Following by that, a 6 x 10^5 of cell density for each stable cell lines were counted and seeded in T-25 flasks. MEL cells are suspension cell lines and do not need a specific surface to adhere on. Therefore, T-25 flasks were used to seed the MEL stable cells for the capacity and easy handling reasons. Two T-25 flasks were seeded for each stable cell line, with one flask was designated for the non-induced condition and the other flask was for DMSO induced condition. After seeding cells for all the stable cell lines, the T-25 flasks were incubated at 37 ºC for 24 hours.

After 24 hours of incubation, MEL stable cell lines in T-25 flasks were checked again under fluorescence microscope to ensure the condition of cells was appropriate for the induction. Subsequently, 100 µL of DMSO (2% vol/vol) was added into the flasks designated for induced condition. The flasks were mixed gently by rocking back and forth. Following by that, the cells were incubated in CO_2 incubator at 37 ºC for 5 consecutive days. At the 3th day of induction, fresh medium with fresh DMSO were replaced and the DMSO induction was then carried on till the 5^{th} day. The induced MEL stable cells were then transferred from the T-25 flasks into 15 mL Falcon tubes and spun down at 300 rpm. Then, supernatants were discarded and pellets in the tube were lysed immediately by adding TRIzol for RNA extraction.

2.6 RNA extraction

Total RNA was isolated from the transfected HeLa cells by using TRIzol reagent. The cells were lysed directly in the 6-well plate by adding 1mL of TRIzol reagent. Cell lysates were then passed through a pipette for several times and incubated for 5 minutes at 15°C to 30°C to permit the complete dissociation of nucleoprotein complexes.

Subsequently, 0.2 mL of chloroform was added into cell lysates and the mixture was shaked vigorously by hand for 15 seconds, followed by an incubation of 2 to 3 minutes at 15°C to 30°C. The samples were then centrifuged at 12,000g for 15 minutes at 2°C to 8°C. Following centrifugation, the mixture was separated into a lower red phenol-choloroform phase, an interphase and a colorless upper aqueous phase. The colorless aqueous phase, which contains RNA, was transferred into a fresh tube.

RNA was then precipitated from the aqueous phase by adding 0.5 mL of isopropylalcohol. The mixture was incubated at 15°C to 30°C for 10 minutes and then centrifuged at 12,000g for 10 minutes at 2°C to 8°C. After RNA precipitated, a visible gel-like pellet was observed at the side and bottom of the tube.

Supernatant was removed from the tube and the RNA pellet was washed with 1 mL of ethanol by centrifuging at 7,500g for 5 minutes at 2° to 8°C. After the washing step, supernatant was removed and RNA pellet was briefly air-dried. RNA pellet was then dissolved in RNAse-free water and was incubated at 55°C to 60°C for

10 minutes. The concentration of RNA was measured using NanoDrop at the absorbance of 260 nm. RNA samples were stored at -86°C until use.

2.7 RNA purification

Before the synthesis of first strand cDNA, each RNA sample was purified using TURBO DNA-free kit (Ambion), to remove contaminating DNA during RNA extraction. From the total RNA that was extracted, an amount of 2.8 µg RNA was used in the purification reaction, which was diluted in 25 µl of DEPC water. 3 µl of 10X TURBO DNase buffer and 2 µl of TURBO DNase were added into the diluted RNA. The reaction mixture was then mixed gently and incubated at 37°C for 1 hour. After the incubation step, 6 µl of DNase inactivation reagent was added and the reaction mixture was incubated at room temperature for 2 minutes. Subsequently, the reaction mixture was spun down at 10,000g for 1.5 minutes. After centrifuging, the supernatant which contained the purified RNA was transferred into a fresh tube and was kept at -86°C. The quality of purified RNA was checked by running gel electrophoresis.

2.8 cDNA synthesis

The first strand of cDNA was synthesized using Taqman Reverse Transcription Reagents (Applied Biosystems) with oligo $d(T)_{16}$ as the primer. The 20 µl reverse transcription reaction was carried out by using 280 ng of RNA treated by TURBO DNase I as the template, with 1 x RT buffer, 5.5 mM $MgCl_2$, 500 µM dNTP mixtures, 2.5 µM oligo $d(T)_{16}$, 0.4 U/µl RNase inhibitor and 1.25 U/µl MultiScribe™ Reverse Transcriptase. The reaction was incubated at 25°C for 10 minutes for the binding of oligo $d(T)_{16}$ primer with purified RNA template, followed by 48ºC for 1

hour of reverse transcription process and finally 95ºC for 5 minutes of reverse transcriptase inactivation.

2.9 Real Time PCR

There are two types of quantitative real time PCR- absolute and relative. In this study, the relative quantification approach was used to analyze the gene expression pattern of different HBB mutations. This was done by comparing the change in expression of HBB gene relative to EGFP gene which served as the calibrator.

The primers used in the real time PCR analysis of β-globin gene and EGFP gene were designed using Primer Express 2.0 (Table 2). All the expression studies for the three experiments were the same, in order to create a fair platform for comparison. The real time PCR reactions were performed in an ABI Prism 7000 Sequence Detection System by using the SYBR Green PCR master mix (Applied Biosystems). A master mix of 2 x SYBR Green PCR master mix and 300 nM of both forward and reverse primers were prepared before being transferred to a optical 96-well reaction plate (Applied Biosystems). A final reaction volume of 25 µl was prepared after adding a total of 6 ng of cDNA template to the master mix. The thermal cycling conditions started with a 10 minutes of incubation at 95ºC to activate the AmpliTaq Gold DNA polymerase in the 2 x SYBR Green PCR master mix. Subsequently, a denaturation step was carried out at 95ºC for 15 seconds and was followed by the annealing and extension steps at 60ºC for 1 minute. The denaturation, annealing and extension steps were then repeated for a total of 40 cycles. For real time PCR reaction, both β-globin and EGFP genes were amplified three times from each sample.

Table 2. HBB and EGFP primers for real time PCR

Name	Sequence 5' to 3'	Length (bp)	Amplicon size
Exon2-F1b (HBB)	TCA AGG GCA CCT TTG CCA CAC	21	86 bp
Exon2/3-Rb (HBB)	CAC GTT GCC CAG GAG CCT G	19	
EGFP mRNA-280F	CTC GAT GTT GTG GCG GAT CT	20	71 bp
EGFP mRNA-350R	ACG TCT ATA TCA TGG CCG ACA AG	23	

After the real time PCR reaction, a dissociation curve analysis was performed by increasing the temperature from 65ºC to 95ºC. This is to differentiate the desired PCR product from primer dimer and also to detect if there is any non-specific product that was being amplified. Subsequently, all the amplified PCR products were run through a 1% agarose gel to check for the desired amplicon size, to further confirm the correct amplification for the real time PCR reaction.

As mentioned earlier, relative quantification approach was used in this project to study the gene expression pattern of different HBB mutations. All the constructs that were used in this project contain both HBB and EGFP genes, except for HeLa Tet-Off experiment, which the co-transfaction was carried out. Hence, the comparative C_T method (ΔC_T) was used to relatively quantitate the β-globin gene mRNA. As for the EGFP gene, it was used as the endogenous control that functions to normalize the differences in transfection efficiency and the amount of template used in the real time PCR reaction. Besides EGFP gene serving as the endogenous control, the total RNA that were extracted from transfected cells were included in the same PCR reaction plate to serve as a negative control. The calculation for the

quantitation involve firstly, finding the difference (ΔC_T) between the C_T values of the target gene (HBB) and the reference gene (EGFP).

$$\mathbf{\Delta C_T = C_T\ (HBB) - C_T\ (EGFP)}$$

This value was calculated for every sample that was to be quantitated. Subsequently, the differences between each sample's ΔC_T will be compared against the ΔC_T of the calibrator, which is the wild type, to generate the value of $\Delta\Delta C_T$.

$$\mathbf{\Delta\Delta C_T = \Delta C_{T\ mutant} - \Delta C_{T\ wild\ type}}$$

The final quantitation step was to transform these values to absolute values. According to the Applied Biosystems guidelines, the PCR efficiency for an amplicon smaller than 150 bp, is close to 100%. Therefore, the amount of target (HBB), normalized to the endogenous control (EGFP) and relative to the calibrator (wild type), is formulated by:

$$\mathbf{Comparative\ expression\ level = 2^{-\Delta\Delta C_T}}$$

ABI relative quantification software was used to automatically calculate the comparative expression level of the β-globin gene from the raw data (C_T values) generated from the real time PCR relative quantification plate documents. The results were exported into Excel files to calculate the average and the standard deviation from three independent transfection repeats for each construct.

3.0 Results

In order to determine the relationship between the position of the PTCs and NMD in β-thalassemia, three PTCs caused by naturally occurring β-thalassemia nonsense or frameshift mutations at different exons were analyzed in this study, namely the codon 41/42, codon 121 and IVS 2-654 mutations. Codon 41/42 is a mutation that occurs in exon 2 of β-globin gene, characterized by a deletion of four nucleotides (-TTCT). The frameshift mutation has caused the presence of PTC at codon 60/61. As for codon 121 mutation, it occurs when there is a substitution of G to T at the position of codon 121 (GAA to TAA), resulting a PTC at codon 121. This nonsense mutation occurs in exon 3 and is known to be a dominantly inheritance mutation. The third mutation, IVS 2-654, occurs when there is a C to T substitution at the position of 654 in intron 2 region, of β-globin gene. This creates a new cryptic splicing site and results in incorporation of extra 73 nucleotides of intron 2 into the open reading frame. Interestingly, this will cause the presence of a PTC at codon 121 too. (Figure 4).

3.1 Steady state gene expression level of PTC-containing transcripts in HeLa cells

Before elucidating the correlation between NMD and position of PTCs in β-globin gene, the steady state gene expression level for codon 41/42, codon 121 and IVS 2-654 were carried out, in order to confirm the initial transcript level of each mutation.

Constructs of either the wild type or mutant (codon 41/42, codon 121 or IVS 2-654) human β-globin genes (pHBB-EGFP) are driven by the β-globin gene

promoter to prevent the possibility of any other promoter effect on the mRNA accumulation. The EGFP gene driven by the CMV promoter was inserted into the same construct containing the β-globin gene to serve as a control for transfection efficiency. The presence of both target gene (β-globin) and reference gene (EGFP) in one construct will ensure the precise normalization against the EGFP expression, hence allowing accurate comparisons between different expressions of varying transfection efficiencies. In addition, EGFP expression will allow visual determination and confirmation of transfection efficiency prior to the downstream experiments. Total RNAs were extracted from cells and the relative gene expression levels of PTC-containing β-globin genes were compared with wild type β-globin gene in transiently transfected HeLa cells after normalizing against EGFP expression.

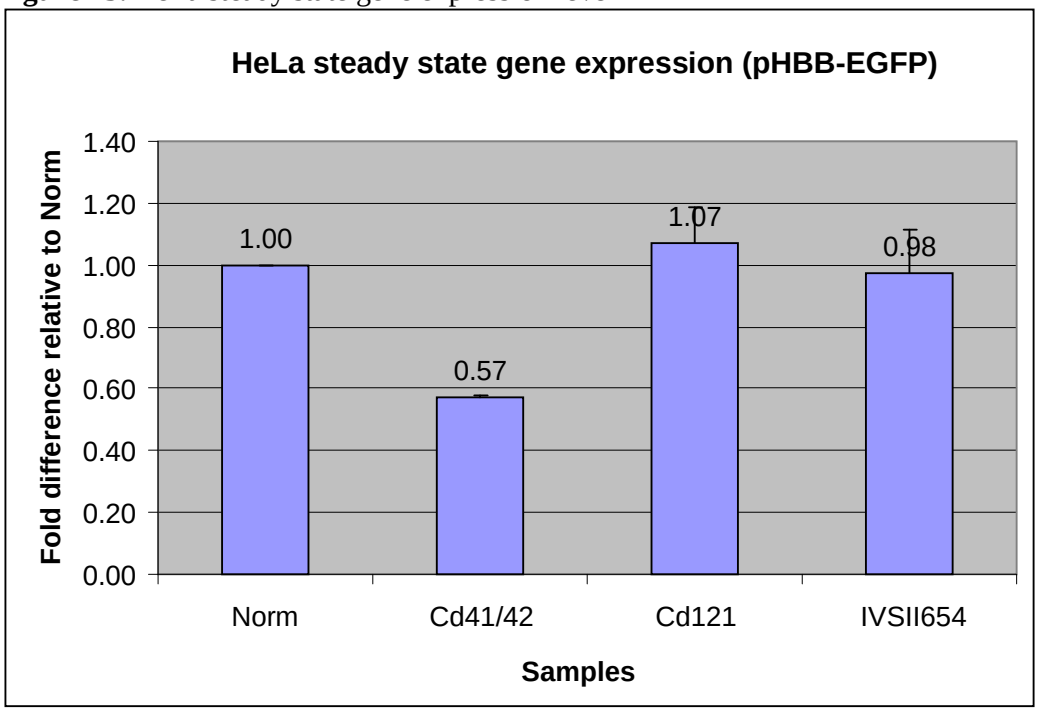
Table 3 shows the summarized results from 3 independent transfection experiments after normalization against EGFP expression. The graph of relative gene expression level for each mutation was plotted as compared to the wild type β-globin (Norm), which serves as the calibrator, with the constant reading of 1. (Figure 15). The relative gene expression level of codon 41/42 was averaged at 57% of wild type transcript level; whereas mutant transcripts for both codon 121 and IVS 2-654 were 98% and 107% respectively, of wild type transcript level. From previous studies, it has been shown that codon 41/42 mutation in exon 2 is susceptible to NMD. Hence, it is expected that the steady state mutant transcript level for codon 41/42 would be reduced. On the other hand, codon 121, containing a nonsense mutation that is located in exon 3 of β-globin gene, is expected to escape NMD. Hence, this has resulted in an accumulation of mutant transcripts, which was slightly higher than wild type transcript level, at 107% of wild type transcript. Although it has been reported

that IVS 2-654 shows a recessive inheritance phenotype, it is very interesting to note that mutant transcripts of IVS 2-654 was observed to accumulate to a level that is comparable to the wild type transcripts (98% of the wild type transcript levels).

Table 3. Summarized results of HeLa steady state gene expression study

	Set 1	Set 2	Set 3	Average	SD
Norm	1	1	1	1.00	0.00
Cd41/42	0.575	0.576	0.563	0.57	0.01
Cd121	0.925	1.218	1.063	1.07	0.12
IVSII654	1.154	0.818	0.955	0.98	0.14

Figure 15. HeLa steady state gene expression level



3.2 mRNA decay rate for PTC-containing transcript in HeLa Tet-Off system

After determining the steady state gene expression levels of PTC-containing construct, it was observed that codon 121 and IVS 2-654 mutations apparently avoided the full impact of NMD. In contrast, codon 41/42 is susceptible to NMD, hence there is reduction in the mutant transcript level. In order to further confirm this, the mRNA decay rate of the three PTCs in β-globin gene as compared to wild type was studied using the Tet-Off system. Each β-globin gene (wild type and mutants) was cloned behind a tet-controlled promoter and trasnfected into a HeLa cell line that stably expresses the tet-transcriptional transactivator (HeLa/tTA cells). A pCMV-EGFP construct was co-transfected to serve as a control to ensure equal transfection efficiency and to show equal amplification efficiency in the real time assays. The cells were transcriptionally pulsed for 4 hours, 8 hours, 12 hours and 24 hours, by adding doxycycline and the corresponding mRNA decay rate was subsequently established. If both codon 121 and IVS 2-654 are able to evade the NMD mechanism, the mRNA decay rate of both these mutations should be comparable to wild type transcript, even when the transcription was being blocked by doxycycline.

Figure 16 shows the summarized gene expression level for each of the cells carrying either the wild type or mutant β-globin genes in HeLa Tet-Off system. The summarized results were from three independent transfection experiments after normalizing against EGFP expression. Figure 17 shows a compilation of the entire mRNA decay rate study for wild type, codon 41/42, codon 121 and IVS 2-654, from four independent experiments, with mRNA levels plotted against time post-Doxycycline treatment. Both codon 121 and IVS 2-654 showed a similar mRNA decay pattern when compared to wild type β-globin. After treating with Doxycycline

for 4 hours, the mRNA levels for codon 121 and IVS 2-654 dropped to 47% and 57% respectively. For codon 41/42, the mRNA level has dropped to 24%. As for wild type β-globin gene, mRNA levels were reduced to 51%. In short, the mRNA decay rate for both codon 121 and IVS 2-654 when compared to wild type β-globin was quite similar. This can be visualized clearly by referring to Figure 17. These results are indeed consistent with the steady state gene expression levels of codon 121 and IVS 2-654, in which both mutations possessed a similar transcript level in comparison to wild type β-globin gene. As for codon 41/42, gene expression levels are comparable in both steady state gene expression and Tet-Off studies.

Figure 16. Summarized gene expression level for wild type, codon 41/42, codon 121 and IVS 2-654 in HeLa Tet-Off system

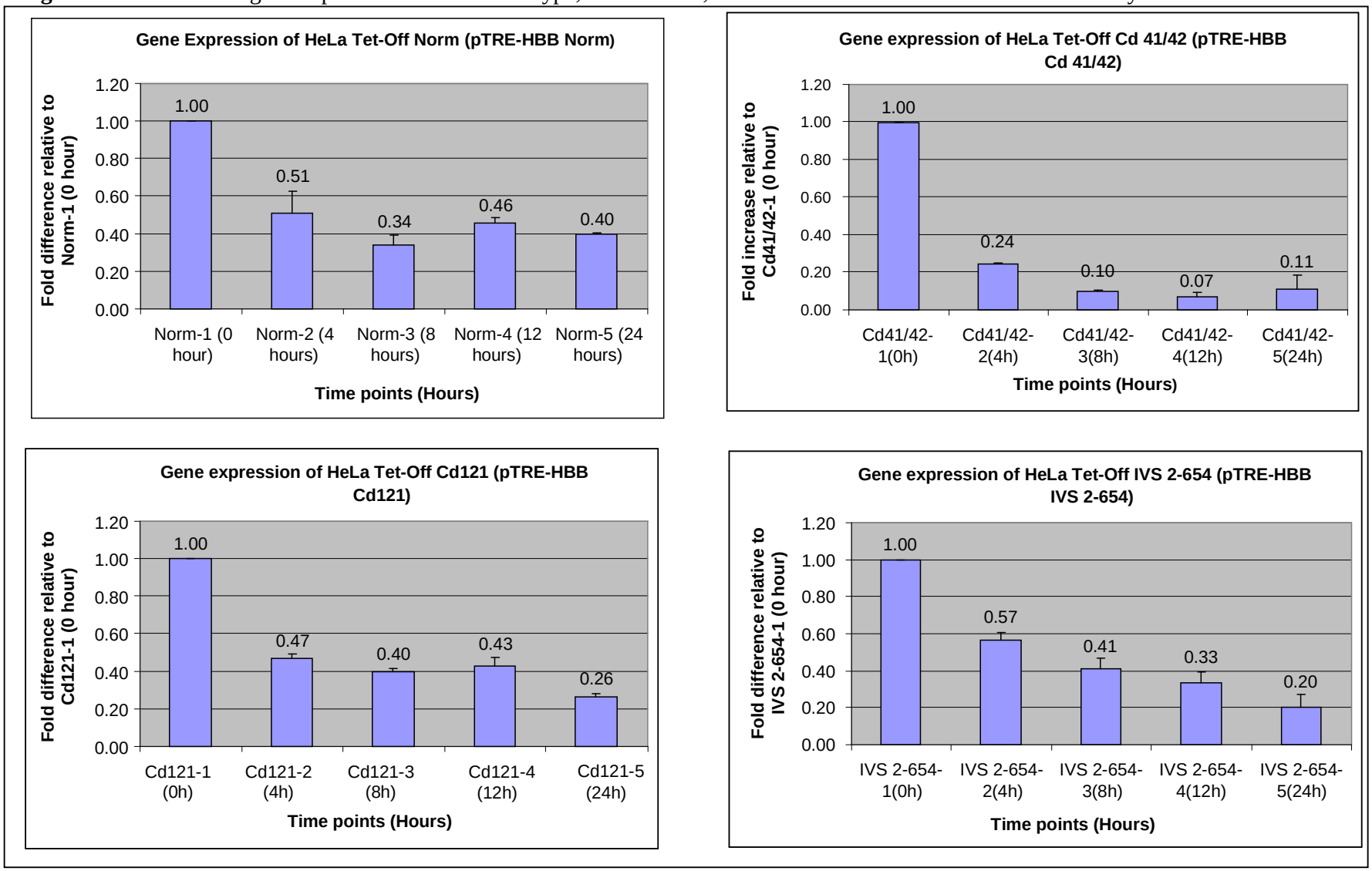
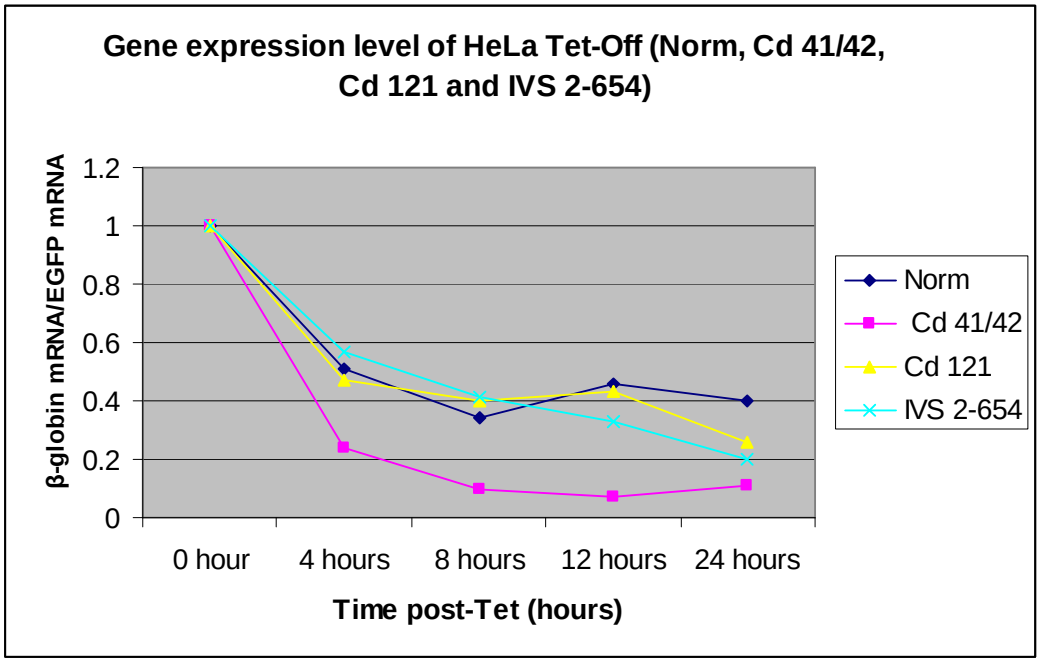


Figure 17. mRNA decay rate for wild type, codon 41/42, codon 121 and IVS 2-654

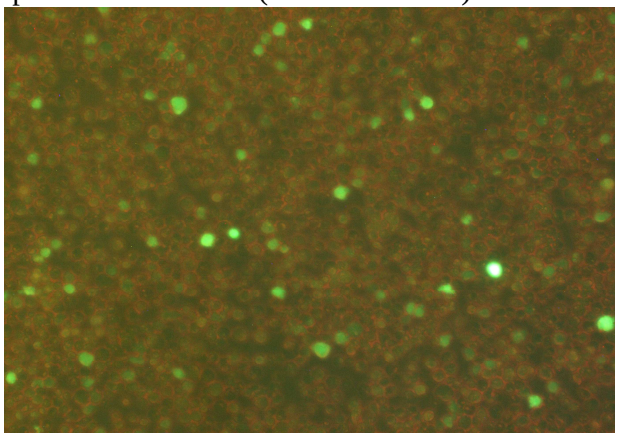


3.3 Gene expression of PTC-containing transcripts in MEL cells (With DMSO/ Without DMSO)

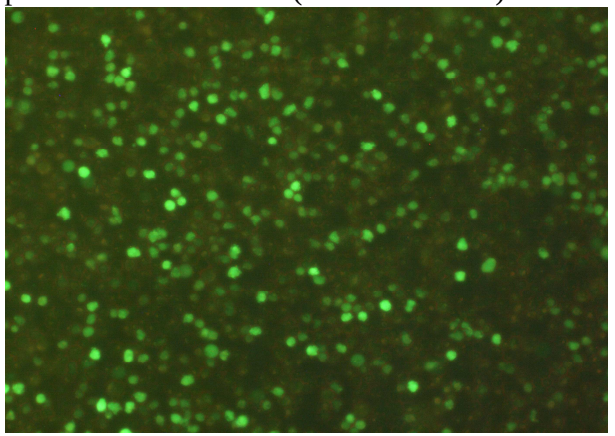
Both the steady state gene expression and mRNA decay rate studies were carried out through transient transfection in HeLa cells. However, although HeLa cell lines are of human origin, they are cervical epithelial cells and hence, could not emulate the native erythroid environment of erythroid cells. Therefore, MEL (mouse erythroleukemia) cells were used to create stable cell lines containing either the wild type or mutant β-globin genes. Figure 18a and 18b show the different MEL stable cell lines that were created to stably express the different pHBB-EGFP constructs. These pHBB-EGFP constructs were also used in the steady state gene expression study.

Figure 18a. MEL stable cell lines (Without and With 2% DMSO)- Norm and codon 41/42

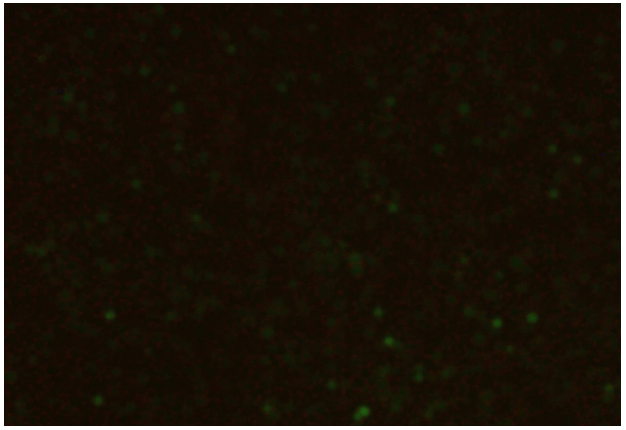
pHBB-EGFP Norm (Without DMSO)



pHBB-EGFP Cd 41/42 (Without DMSO)



pHBB-EGFP Norm (Treated with 2% DMSO for 5 days)



pHBB-EGFP Cd 41/42 (Treated with 2% DMSO for 5 days)

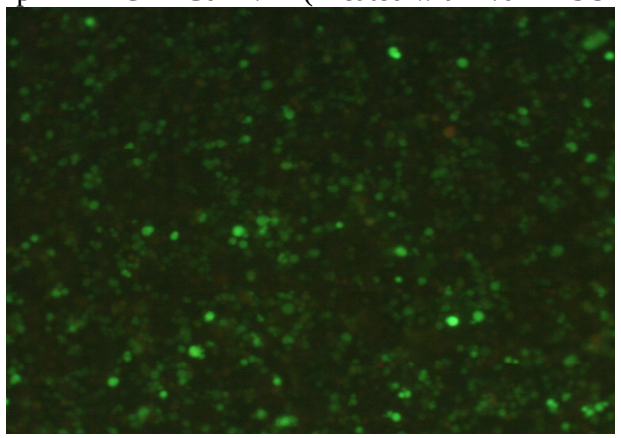
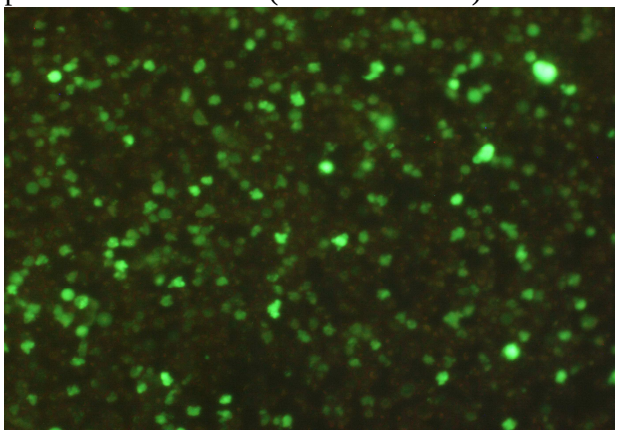
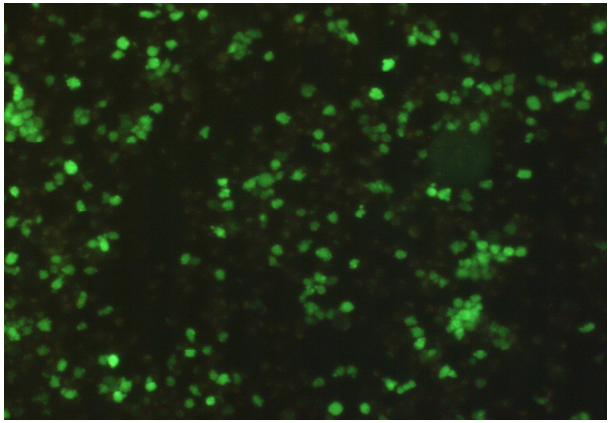


Figure 18b. MEL stable cell lines (Without and With 2% DMSO)- codon 121 and IVS 2-654

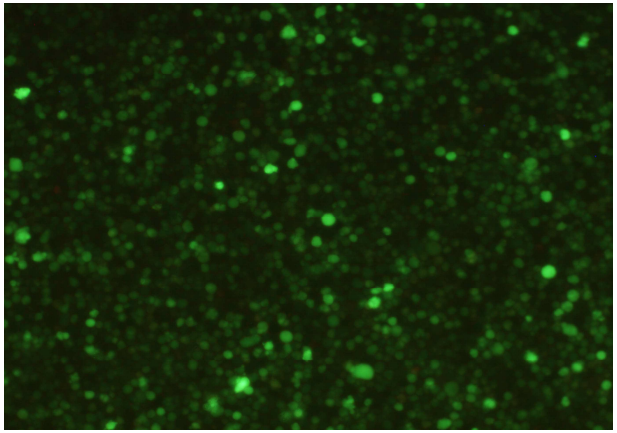
pHBB-EGFP Cd 121 (Without DMSO)



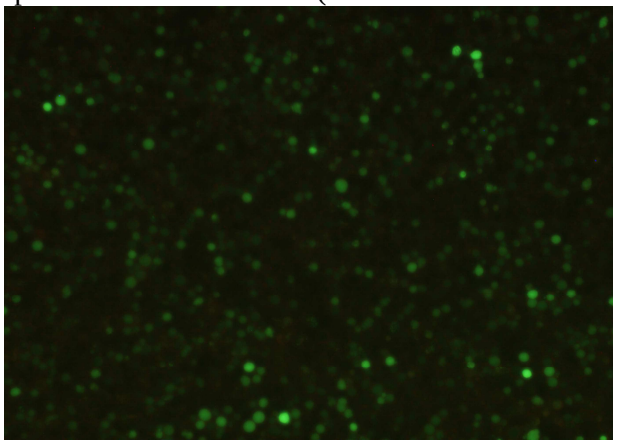
pHBB-EGFP IVS 2-654 (Without DMSO)



pHBB-EGFP Cd 121 (Treated with 2% DMSO for 5 days)



pHBB-EGFP IVS 2-654 (Treated with 2% DMSO for 5 days)



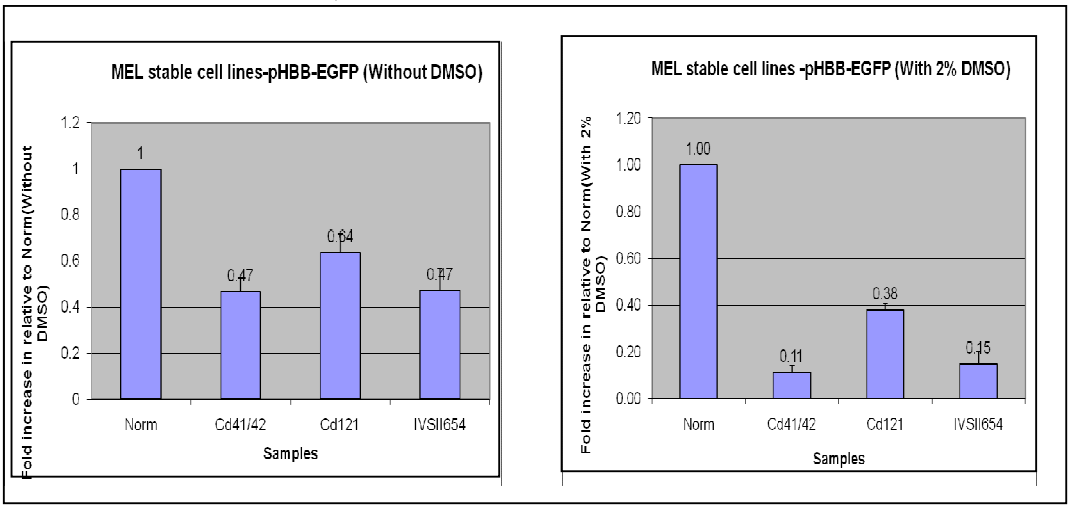
In order to create an environment that mimic matured erythroblasts, these MEL stable cell lines were terminally differentiated by adding 2% of DMSO for 5 days. Total RNAs were extracted from cells and real time PCR was performed to quantify the levels of β-globin mRNA accumulation before and after the induction of MEL cell erythroid differentiation.

Figure 18a and Figure 18b showed the image of MEL stable cell lines of before and after erythroid differentiation by the addition of 2% DMSO. Due to the toxicity of DMSO, it can be observed that the incubation of MEL stable cell lines with DMSO for 5 days resulted in a certain degree of cell death. Hence, the fluorescence intensity of the stable cell lines was reduced compared to the the intensity before DMSO induction, especially for MEL stable cell line carrying the pHBB-EGFP Norm construct.

Figure 19 showed the gene expression level for MEL stable lines that were without and with DMSO induction. Expression of each mutant allele was compared to wild type gene, after normalizing against the EGFP expression. Intriguingly, the steady state gene expression levels in MEL stable cell lines (without DMSO induction), showed a different result from what the levels observed in HeLa cells. Codon 121 and IVS 2-654 mutant transcript were averaged at 64% and 47%, respectively, compared to wild type β-globin gene in MEL cells. The mutant mRNA transcripts accumulated for both mutations were significantly lower, comparing to the levels observed in HeLa cells. As for codon 41/42 mutation, the accumulated mutant transcripts in both MEL and Hela cells were quite consistent, averaging at 47% in comparison to wild type β-globin.

Interestingly, after the induction of erythroid differentiation, the gene expression levels for all the mutations, codon 41/42, codon 121 and IVS 2-654 in MEL cells were relatively low as compared to the corresponding wild type β-globin gene. For codon 41/42 mutation, the mutant transcript was expressed at about 11% of the induced wild type β-globin gene expression. This was significantly lower compared to the non-induced mutant transcript of codon 41/42, which was about 47% of wild type β-globin gene. For codon 121 and IVS 2-654 in differentiated MEL cells, both mutations only averaged 38% and 15% of wild type β-globin gene, respectively. This data indicates that there might be another type of surveillance mechanism like NMD or factors being triggered to target on mutant transcripts in erythroid cells.

Figure 19. Gene expression level of MEL stable cell lines (before and after differentiation)



4.0 Discussion

The mechanism of NMD has been studied for almost 20 years (Holbrook JA 2004). In mammalian cells, it was first discovered during the studies of β^0-thalassemias caused by PTCs (Maquat 2004). Since then, many studies have been carried out with the aim to elucidate the mechanism of NMD, especially in β-thalassemia. Several studies have been focusing on drawing a boundary for the location of PTCs that manage to trigger NMD and hence reduce the abundance of human β-globin mRNA. A "3' boundary rule" has been established, that all PTCs located more than 50-55 nucleotides from the final exon-exon junction will elicit NMD (Nagy and Maquat 1998; Zhang, Sun et al. 1998). However, a number of exceptions to the 50-55 nucleotide boundary rule have been reported. One of these cases of NMD-resistance include human β-globin transcripts with PTCs located at 5' of intron 1 are shown to be resistant to NMD (Romao, Inacio et al. 2000). This has suggested that the "3' boundary rule" is not commonly applicable to all the PTCs that are found on human β-globin gene and it is possible that there could be other boundaries and regions that are able or unable to elicit NMD mechanism.

Hence, to further elucidate the relationship between locations of PTC and with NMD in β-thalassemia, three naturally occurring mutations (at codon 41/42, codon 121 and IVS 2-654) that caused PTCs in exon 2 and exon 3 of human β-globin gene were analyzed. The levels of PTC-containing transcripts were compared with wild type transcript to study the effect of NMD in vitro.

The steady state gene expression levels of these three mutations in a HeLa transient transfection experiment have shown to comply with the "50-55 nucleotides" boundary rule. For these three mutations, only codon 41/42 is located in exon 2 of β-globin gene, with PTC at codon 60/61 and both codons 121 and IVS 2-654 mutations are located in exon 3 and intron 2 respectively, with both PTCs at codon 121. As expected, our data showed that 57% of codon 41/42 mutant transcript as compared to wild type β-globin gene was obtained and this has suggested NMD was being triggered and resulted in the halved mutant transcript level as compared to wild type. As for codon 121 and IVS 2-654, with both PTCs located at codon 121, exon 3 of human β-globin gene, high level of mutant transcripts were obtained, which are slightly higher than the level of wild type transcript. It is expected that the dominantly inherited codon 121 mutation will have accumulated a substantial amount of aberrant transcripts but not for the recessively inheritance IVS 2-654 mutation. This is an interesting observation for IVS 2-654 because it has been well documented that in cases with PTCs in third exon, NMD is not elicited and the accumulation of aberrant transcripts will lead to an apparent dominant phenotype (Thein, Hesketh et al. 1990; Kazazian, Dowling et al. 1992). However, it has to be taken into consideration that this steady state gene expression study was carried out in a HeLa cell, which is not an erythroid cell line, hence giving some intriguing results that were not reported in previous studies. Generally, this result of steady state gene expression showed the effect of PTCs position in NMD for the three mutations, suggesting that both codon 121 and IVS 2-654 manage to evade NMD, due to the location of their PTCs Hence, we can conclude that in HeLa cells, the 50-55 nucleotide boundary rule is shown to be applicable to PTCs, which are specifically located in exon 2 and exon 3, and this is consistent with most of the reported studies (Takeshita, Forget et al. 1984; Baserga

and Benz 1988; Lim, Sigmund et al. 1992; Hall and Thein 1994; Ho, Wickramasinghe et al. 1997; Romao, Inacio et al. 2000)

In order to further confirm the steady state gene expression levels that we have observed, a transcriptional pulse chase (Tet-Off) strategy was used to study the mRNA decay pattern of the three mutations as compared to wild type human β-globin gene. As expected, the results from the Tet-Off system appeared to recapitulate our observations in normal HeLa cells. For codon 41/42, a prototype NMD-sensitive nonsense mutation, the mRNA decay rate was almost two times faster compared to wild type β-globin. This has again confirmed that codon 41/42 mutation triggered NMD and accumulated much lower level of mutant transcript at initial stage. Hence, this explained the fast decay rate of codon 41/42 in this transcriptional pulse chase study. Concordant with the HeLa steady state gene expression results, both codons 121 and IVS 2-654 portrayed a similar pattern of mutant transcript decay rate when compared to wild type, indicating the presence of high level of mutant transcript at 0 hour time point. In short, in a non-erythroid cells environment such as HeLa cells, PTC located in exon 2 will elicit NMD and hence producing a much lower mutant transcript as compared to wild type β-globin gene. As for PTCs located in exon 3, they are resistant to NMD and hence will accumulate substantial amount of mutant transcripts.

HeLa cells were derived from cervical epithelial cells, a non-erythroid cell line. Although transfection reaction is more feasible in HeLa cells, gene expression study utilizing HeLa cells is not able to emulate the actual *in vivo* environment of eythroid cells. Hence, we have established mouse erythroleukemia cells (MEL) stable cell

lines, carrying either the wild type or mutant human β-globin genes. In addition to using these MEL stable cells lines for gene expression studies, erythroid differentiation was also induced in these cell lines with the addition of DMSO to study the expression levels of both wild type and mutant β-globin genes in a cell type that closely mimic an *in vivo* environment. Besides that, this was carried out in order to elucidate the actual gene expression level of IVS 2-654 in both HeLa and MEL cell environment.

In the MEL gene expression studies, all the three mutations, codon 41/42, codon 121 and IVS 2-654 have shown a sharp drop in accumulation of mutant transcript level after DMSO induction. For codon 41/42, the low expression level observed is as expected as this mutation is known to elicit NMD, and hence, only a small amount of mRNA mutant transcript was detected. An interesting observation from this study is the expression levels of codons 121 and IVS 2-654, both only averaged 38% and 15% of wild type β-globin, respectively. Codon 121 was the first nonsense codon mutation in exon 3 and is known to result in a dominantly inherited form of β-thalassemia (Stamatoyannopoulos, Woodson et al. 1974; Fei, Stoming et al. 1989). According to the 50-55 nucleotide boundary rule, codon 121 will escape NMD leading to the accumulation of a significant amount of mRNA transcript. *In vivo* evidences have shown that the amount of codon 121 mutant transcript was at least equal to the normal β-globin transcript in patients with codon 121 mutation (Hall and Thein 1994; Ho, Wickramasinghe et al. 1997). However, codon 121 mutant transcript level that we have obtained is lower than the reported studies. The discrepancy between our results and the two published reports may be attributed to the different sources of starting material that was being used. In our study, total RNA

was extracted from differentiated MEL (mouse erythroleukemia) stable cell lines. As for the two published reports, studies were carried out using reticulocytes from patients and subsequently assayed by conventional RT-PCR. In the first place, it is apparent that although we have used a differentiated erythroid cell as our study model, the NMD mechanism in mouse might not be the fully comparable to human erythroid cells. However, in quantitative terms, our results still showed that codon 121 nonsense mutation leads to measurable mRNA accumulation.

Apart from codon 121, IVS 2-654 also showed a relatively low mutant transcript level after DMSO induction, averaging only 15% of wild type β-globin transcript level. This is also an interesting phenomenon in this study. IVS 2-654, with a mutation from C to T in the intervening sequence 2 position 654, results in the incorporation of 73 extra nucleotides of intron 2 into the aberrantly spliced mRNA transcript. This mutation results in an abnormally spliced mRNA that, if translated, would lead to a PTC at codon 121. According to the 50-55 nucleotide boundary rule, the location of codon 121 in exon 3 will cause IVS 2-654 to evade NMD and in turn should produce a substantial amount of mutant transcript. However, interestingly, the majority of cases that have been described showed an asymptomatic phenotype but not the pattern of dominant β-thalassemia. An *in vivo* study of an IVS 2-654 asymptomatic case has revealed that large amounts of aberrantly spliced mRNA from the IVS 2-654 allele were detectable only in the early erythroblasts stage. It was also reported that a large decrease in the amount of the abnormally spliced IVS 2-654 mRNA was observed during the maturation of erythroblasts to reticulocytes, suggesting the instability of the mutant transcript (Ho, Hall et al. 1998). Clearly, this *in vivo* study of the IVS 2-654 mutation is in supported of what we have observed in

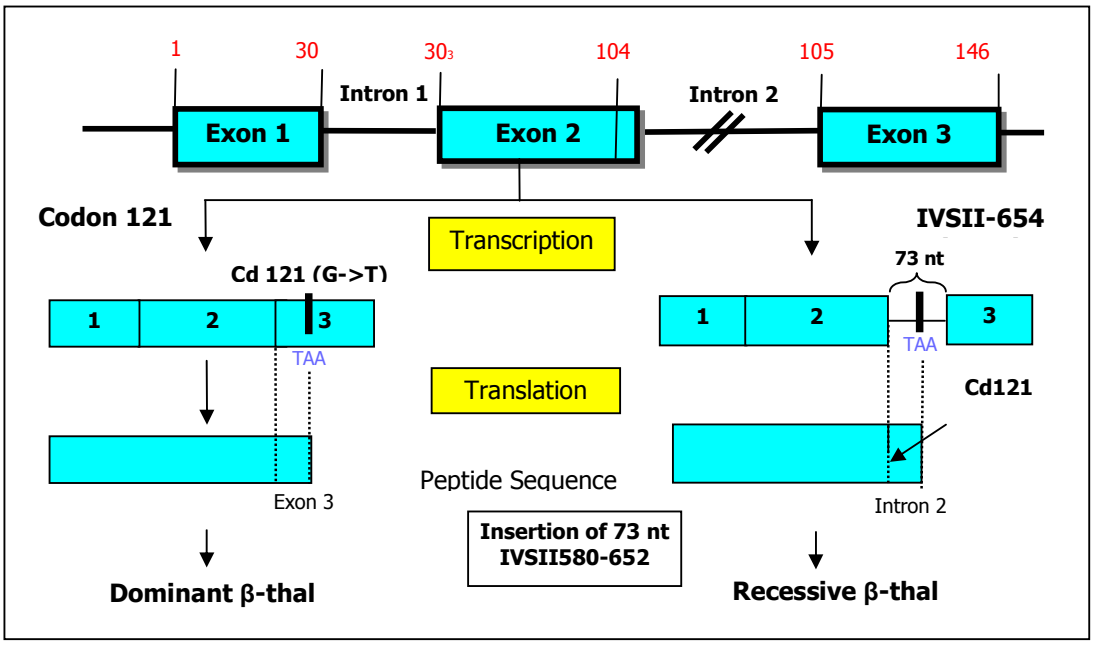
the *in vitro* MEL stable cell line system, as we have also shown a reduction of IVS 2-654 mutant transcripts after the erythroid differentiation. On the other hand, this has shown that the steady state gene expression level of IVS 2-654 that we have obtained in HeLa cells might not reflect the actual gene expression level of IVS 2-654 in an in vivo environment. A possible explanation for this observation might be the presence of erythroid transcription factors in MEL cells, which are not found in HeLa, the non-erythroid cell line. The eyrthroid transcription factors such as Erythroid Kruppel Like Factor (EKLF), GATA-1 and NF-E2 are the main players in regulating β-globin expression (Palstra 2009). Hence, without the interaction of these erythroid transcription factors with β-globin gene, the actual gene expression level of IVS 2-654 was not faithfully reproduced in HeLa system.

From our observation, the 50-55 nucleotide rule of NMD does not apply on IVS 2-654 mutation. Unlike the codon 121 mutation, IVS 2-654 averaged a relatively low mutant transcript level compared to wild type β-globin. Hence, it is interesting to note that although both codons 121 and IVS 2-654 possess PTCs at codon 121 in exon 3, these two mutations have shown to accumulate different amounts of mutant transcripts in our *in vitro* system (Figure 19). Another interesting fact is that codon 121 is a dominant inheritance mutation whereas IVS 2-654 is of recessive inheritance. However, the only difference between these two mutations is that the mutant IVS 2-654 transcript that is aberrantly spliced, contain an extra 73 nucleotides from intron 2 sequence (Figure 20). It is therefore possible that this extra intronic sequence in the mutant IVS 2-654 transcript contributes to its low expression levels. In support of this, it has been reported that introns do play a role in regulating the rate of mRNA decay (Zhang, Sun et al. 1998; Zhao and Hamilton 2007) and it was shown that in the

event of intron retention in transcripts, small introns will undergo strong selective pressure to encode PTCs (Jaillon, Bouhouche et al. 2008). In other words, if introns are not spliced out in transcripts, premature translation termination will take place. Although that study was focused on *Paramecium tetraurelia*, this finding was also observed among the short introns of plants, fungi and animals. Hence, it is possible that the large decrease of accumulated IVS 2-654 aberrant mRNA after erythroid differentiation was due to the effect of NMD but not the instability of mutant transcripts. Apparently, the extra 73 nucleotides of the intron 2 sequence that are incorporated with mRNA transcript is the key factor that elicits NMD. Although it is still unclear how short intronic sequences triggers NMD, this is something novel that can be further studied.

In conclusion, we managed to show that the 50-55 nucletide boundary rule is applicable to codon 41/42, codon 121 and IVS 2 654 in HeLa cells. It is apparent that codon 41/42 which located in exon 2, was subjected to NMD. As for codon 121 and IVS 2-654 in exon 3 and intron 2, they have managed to escape from NMD hence resulted in high gene expression level. On the other hand, in MEL cells in vitro system, we managed to faithfully reproduce the level of the initial mutant transcript and its instability during erythroid cell maturation for codon 121 and IVS 2-654, with the expression levels in vivo. However, it is clear that the 50-55 nucleotide boundary rule does not apply on IVS 2-654. We have speculated that the short 73 nucleotide of intronic sequence in IVS 2-654 mutant transcript managed to elicit NMD, hence contributed to the low expression level and recessive inheritance phenotype of IVS 2-654 mutation.

Figure 20. Difference in between codon 121 and IVS 2-654 mutation



Bibliography

Adams, J. G., 3rd, M. H. Steinberg and H. H. Kazazian, Jr. (1990). "Isolation and characterization of the translation product of a beta-globin gene nonsense mutation (beta 121 GAA----TAA)." Br J Haematol **75**(4): 561-7.

Baker, K. E. and R. Parker (2004). "Nonsense-mediated mRNA decay: terminating erroneous gene expression." Curr Opin Cell Biol **16**(3): 293-9.

Baserga, S. J. and E. J. Benz, Jr. (1988). "Nonsense mutations in the human beta-globin gene affect mRNA metabolism." Proc Natl Acad Sci U S A **85**(7): 2056-60.

Chan, V., T. K. Chan, F. F. Chebab and D. Todd (1987). "Distribution of beta-thalassemia mutations in south China and their association with haplotypes." Am J Hum Genet **41**(4): 678-85.

Chang, Y. F., J. S. Imam and M. F. Wilkinson (2007). "The nonsense-mediated decay RNA surveillance pathway." Annu Rev Biochem **76**: 51-74.

Deisseroth, A., A. Nienhuis, J. Lawrence, R. Giles, P. Turner and F. H. Ruddle (1978). "Chromosomal localization of human beta globin gene on human chromosome 11 in somatic cell hybrids." Proc Natl Acad Sci U S A **75**(3): 1456-60.

Deisseroth, A., A. Nienhuis, P. Turner, R. Velez, W. F. Anderson, F. Ruddle, J. Lawrence, R. Creagan and R. Kucherlapati (1977). "Localization of the human alpha-globin structural gene to chromosome 16 in somatic cell hybrids by molecular hybridization assay." Cell **12**(1): 205-18.

Deisseroth, A., R. Velez and A. W. Nienhuis (1976). "Hemoglobin synthesis in somatic cell hybrids: independent segregation of the human alpha- and beta-globin genes." Science **191**(4233): 1262-4.

Dimaano, C. and K. S. Ullman (2004). "Nucleocytoplasmic transport: integrating mRNA production and turnover with export through the nuclear pore." Mol Cell Biol **24**(8): 3069-76.

Eleftherious (2003). About thalassemia, Thalassemia Internationl Federation.

Fei, Y. J., T. A. Stoming, A. Kutlar, T. H. Huisman and G. Stamatoyannopoulos (1989). "One form of inclusion body beta-thalassemia is due to a GAA----TAA mutation at codon 121 of the beta chain." Blood **73**(4): 1075-7.

Frischmeyer, P. A. and H. C. Dietz (1999). "Nonsense-mediated mRNA decay in health and disease." Hum Mol Genet **8**(10): 1893-900.

Fritsch, E. F., R. M. Lawn and T. Maniatis (1980). "Molecular cloning and characterization of the human beta-like globin gene cluster." Cell **19**(4): 959-72.

Gusella, J., A. Varsanyi-Breiner, F. T. Kao, C. Jones, T. T. Puck, C. Keys, S. Orkin and D. Housman (1979). "Precise localization of human beta-globin gene complex on chromosome 11." Proc Natl Acad Sci U S A **76**(10): 5239-42.

Hall, G. W. and S. Thein (1994). "Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: a mechanism for the phenotype of dominant beta-thalassemia." Blood **83**(8): 2031-7.

Hardison, R. C., I. Sawada, J. F. Cheng, C. K. Shen and C. W. Schmid (1986). "A previously undetected pseudogene in the human alpha globin gene cluster." Nucleic Acids Res **14**(4): 1903-11.

Harju, S., K. J. McQueen and K. R. Peterson (2002). "Chromatin structure and control of beta-like globin gene switching." Exp Biol Med (Maywood) **227**(9): 683-700.

Higgs, D. R. and D. J. Weatherall (2009). "The alpha thalassaemias." Cell Mol Life Sci **66**(7): 1154-62.

Ho, P. J., G. W. Hall, S. Watt, N. C. West, J. W. Wimperis, W. G. Wood and S. L. Thein (1998). "Unusually severe heterozygous beta-thalassemia: evidence for an interacting gene affecting globin translation." Blood **92**(9): 3428-35.

Ho, P. J., S. N. Wickramasinghe, D. C. Rees, M. J. Lee, A. Eden and S. L. Thein (1997). "Erythroblastic inclusions in dominantly inherited beta thalassemias." Blood **89**(1): 322-8.

Holbrook, J. A., G. Neu-Yilik, M. W. Hentze and A. E. Kulozik (2004). "Nonsense-mediated decay approaches the clinic." Nat Genet **36**(8): 801-8.

Hosoda, N., Y. K. Kim, F. Lejeune and L. E. Maquat (2005). "CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells." Nat Struct Mol Biol **12**(10): 893-901.

Jaillon, O., K. Bouhouche, J. F. Gout, J. M. Aury, B. Noel, B. Saudemont, M. Nowacki, V. Serrano, B. M. Porcel, B. Segurens, A. Le Mouel, G. Lepere, V. Schachter, M. Betermier, J. Cohen, P. Wincker, L. Sperling, L. Duret and E. Meyer (2008). "Translational control of intron splicing in eukaryotes." Nature **451**(7176): 359-62.

Jeffreys, A. J., I. W. Craig and U. Francke (1979). "Localisation of the G gamma-, A gamma-, delta- and beta-globin genes on the short arm of human chromosome 11." Nature **281**(5732): 606-8.

Kazazian, H. H., Jr., C. E. Dowling, R. L. Hurwitz, M. Coleman, A. Stopeck and J. G. Adams, 3rd (1992). "Dominant thalassemia-like phenotypes associated with mutations in exon 3 of the beta-globin gene." Blood **79**(11): 3014-8.

Kinniburgh, A. J., L. E. Maquat, T. Schedl, E. Rachmilewitz and J. Ross (1982). "mRNA-deficient beta o-thalassemia results from a single nucleotide deletion." Nucleic Acids Res **10**(18): 5421-7.

Koeffler, H. P., R. S. Sparkes, H. Stang and T. Mohandas (1981). "Regional assignment of genes for human alpha-globin and phosphoglycollate phosphatase to the short arm of chromosome 16." Proc Natl Acad Sci U S A **78**(11): 7015-8.

Le Hir, H., D. Gatfield, E. Izaurralde and M. J. Moore (2001). "The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay." Embo J **20**(17): 4987-97.

Le Hir, H., E. Izaurralde, L. E. Maquat and M. J. Moore (2000). "The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions." Embo J **19**(24): 6860-9.

Lebo, R. V., A. V. Carrano, K. Burkhart-Schultz, A. M. Dozy, L. C. Yu and Y. W. Kan (1979). "Assignment of human beta-, gamma-, and delta-globin genes to the short arm of chromosome 11 by chromosome sorting and DNA restriction enzyme analysis." Proc Natl Acad Sci U S A **76**(11): 5804-8.

Lejeune, F., X. Li and L. E. Maquat (2003). "Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities." Mol Cell **12**(3): 675-87.

Li, S. and M. F. Wilkinson (1998). "Nonsense surveillance in lymphocytes?" Immunity **8**(2): 135-41.

Lim, S. K., C. D. Sigmund, K. W. Gross and L. E. Maquat (1992). "Nonsense codons in human beta-globin mRNA result in the production of mRNA degradation products." Mol Cell Biol **12**(3): 1149-61.

Lykke-Andersen, J. (2001). "mRNA quality control: Marking the message for life or death." Curr Biol **11**(3): R88-91.
Maquat, L. E. (2002). "Nonsense-mediated mRNA decay." Curr Biol **12**(6): R196-7.
Maquat, L. E. (2004). "Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics." Nat Rev Mol Cell Biol **5**(2): 89-99.
Maquat, L. E. and G. G. Carmichael (2001). "Quality control of mRNA function." Cell **104**(2): 173-6.
Maquat, L. E., A. J. Kinniburgh, E. A. Rachmilewitz and J. Ross (1981). "Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia." Cell **27**(3 Pt 2): 543-53.
Nagy, E. and L. E. Maquat (1998). "A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance." Trends Biochem Sci **23**(6): 198-9.
Naritomi, Y., H. Nakashima, M. Kagimoto, Y. Naito, E. Yokota and T. Imamura (1990). "A common Chinese beta-thalassemia mutation found in a Japanese family." Hum Genet **84**(5): 480-2.
Nicholls, R. D., J. A. Jonasson, J. O. McGee, S. Patil, V. V. Ionasescu, D. J. Weatherall and D. R. Higgs (1987). "High resolution gene mapping of the human alpha globin locus." J Med Genet **24**(1): 39-46.
Orphanides, G. and D. Reinberg (2002). "A unified theory of gene expression." Cell **108**(4): 439-51.
Palstra, R.-J. S. (2009). "Close encounters of the 3C kind: long-range chromatin interactions and transcriptional regulation." BRIEFINGS IN FUNCTIONAL GENOMICS AND PROTEOMICS.
Ribeiro, D. M. and M. F. Sonati (2008). "Regulation of human alpha-globin gene expression and alpha-thalassemia." Genet Mol Res **7**(4): 1045-53.
Romao, L., A. Inacio, S. Santos, M. Avila, P. Faustino, P. Pacheco and J. Lavinha (2000). "Nonsense mutations in the human beta-globin gene lead to unexpected levels of cytoplasmic mRNA accumulation." Blood **96**(8): 2895-901.
Sanders-Haigh, L., W. F. Anderson and U. Francke (1980). "The beta-globin gene is on the short arm of human chromosome 11." Nature **283**(5748): 683-6.
Schechter, A. N. (2008). "Hemoglobin research and the origins of molecular medicine." Blood **112**(10): 3927-38.
Schell, T., A. E. Kulozik and M. W. Hentze (2002). "Integration of splicing, transport and translation to achieve mRNA quality control by the nonsense-mediated decay pathway." Genome Biol **3**(3): REVIEWS1006.
Schrier, S. L. (2002). "Pathophysiology of thalassemia." Curr Opin Hematol **9**(2): 123-6.
Stamatoyannopoulos, G., R. Woodson, T. Papayannopoulou, D. Heywood and S. Kurachi (1974). "Inclusion-body beta-thalassemia trait. A form of beta thalassemia producing clinical manifestations in simple heterozygotes." N Engl J Med **290**(17): 939-43.
Takeshita, K., B. G. Forget, A. Scarpa and E. J. Benz, Jr. (1984). "Intranuclear defect in beta-globin mRNA accumulation due to a premature translation termination codon." Blood **64**(1): 13-22.
Thein, S. L. (1998). "Beta-thalassaemia." Baillieres Clin Haematol **11**(1): 91-126.
Thein, S. L. (1999). "Is it dominantly inherited beta thalassaemia or just a beta-chain variant that is highly unstable?" Br J Haematol **107**(1): 12-21.

Thein, S. L., C. Hesketh, P. Taylor, I. J. Temperley, R. M. Hutchinson, J. M. Old, W. G. Wood, J. B. Clegg and D. J. Weatherall (1990). "Molecular basis for dominantly inherited inclusion body beta-thalassemia." Proc Natl Acad Sci U S A **87**(10): 3924-8.

Thein, S. L., J. M. Old, J. S. Wainscoat, M. Petrou, B. Modell and D. J. Weatherall (1984). "Population and genetic studies suggest a single origin for the Indian deletion beta thalassaemia." Br J Haematol **57**(2): 271-8.

Thermann, R., G. Neu-Yilik, A. Deters, U. Frede, K. Wehr, C. Hagemeier, M. W. Hentze and A. E. Kulozik (1998). "Binary specification of nonsense codons by splicing and cytoplasmic translation." Embo J **17**(12): 3484-94.

Varawalla, N. Y., J. M. Old, R. Sarkar, R. Venkatesan and D. J. Weatherall (1991). "The spectrum of beta-thalassaemia mutations on the Indian subcontinent: the basis for prenatal diagnosis." Br J Haematol **78**(2): 242-7.

Weatherall, D. J. (1998). "Pathophysiology of thalassaemia." Baillieres Clin Haematol **11**(1): 127-46.

Weatherall, D. J. and J. B. Clegg (2001). "Inherited haemoglobin disorders: an increasing global health problem." Bull World Health Organ **79**(8): 704-12.

Weatherall, D. J. and J. B. Clegg (2001). The Thalassemia Syndromes, Oxford.

Weatherall, D. J., G. Stamatoyannopoulous, P. W. Majerus, R. M. Perlmutter and H. Varmus (2001). The molecular basis of blood diseases.

Zhang, J., X. Sun, Y. Qian and L. E. Maquat (1998). "Intron function in the nonsense-mediated decay of beta-globin mRNA: indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm." Rna **4**(7): 801-15.

Zhao, C. and T. Hamilton (2007). "Introns regulate the rate of unstable mRNA decay." J Biol Chem **282**(28): 20230-7.