CHAPTER I

INTRODUCTION

1.1 Hemoglobin

Hemoglobin is the principal protein of the red cell. It is the most abundant protein in human and represents more than 95% of the soluble protein content of the erythrocytes. The major role of the hemoglobin is transportation of oxygen from lung to body tissues. And associated function is the binding of carbon dioxide and proton by deoxyhemoglobin, thereby serving to buffer the blood on the venous side of the circulation. Hemoglobin synthesis requires the coordinated production of heme and globin. The normal hemoglobin molecule contains two types of distinct globin chains; one is designated α-like globin chain which contains 141 amino acid residues. The second globin chain is called β-like globin chain which contains 146 amino acid residues. Three types of hemoglobin are commonly found in normal adults comprising HbA, HbA_2, HbF (Weatherall, 1981). The hemoglobins normally seen in fetal life are HbF and HbA. However, there are three embryonic hemoglobins, Hbs Gower 1 and 2 and Hb Portland. The production of these different hemoglobins is a reflection of a series of physiological adaptations to different oxygen requirements at various stages of development.

1.2 Structure of hemoglobins

All normal human hemoglobins are tetramers consisting of two pairs of unlike globin chains. Adult and fetal hemoglobins have α-globin chains associated with β (HbA, $\alpha_2\beta_2$) or γ-globin chains (HbF, $\alpha_2\gamma_2$), whereas in the embryo ζ-globin chain

combines with γ (Hb Portland, $\zeta_2\gamma_2$) or ε -globin chains (Hb Gower 1, $\zeta_2\varepsilon_2$) and α and ε-globin chains conbine to form Hb Gower 2 ($\alpha_2\varepsilon_2$) (Figure 1.1).

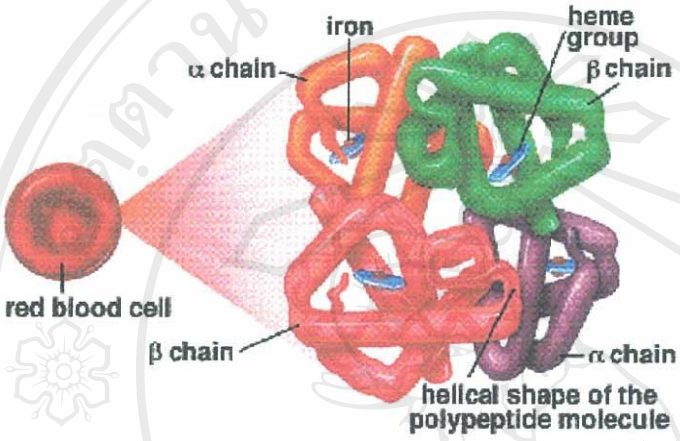

Figure 1.1 Model of hemoglobin A ($\alpha_2\beta_2$), a view of the subunit contacts and the heam pockets. (*http://www.ags.uci.edu/~bcrourke/Blood*)

1.3 Developemental changes of human hemoglobin

The developmental changes in hemoglobin production are brought about by differential activation of the globin genes, which is largely determined at the level of transcription. This series of events is called 'hemoglobin switching'. Human hemoglobin is heterogeneous at all stages of development, beginning with the youngest embryos. In embryos, hemoglobin synthesis is confined to the yolk sac, where Hb Gower 1 ($\zeta_2\varepsilon_2$), Gower 2 ($\alpha_2\varepsilon_2$) and Portland ($\zeta_2\gamma_2$) are produced. At around 7-8 weeks' gestation the liver becomes the major site of erythropoiesis, producing large ennuclated red cells, θ-gene transcripts are detected in yolksac, fetal liver. Very small amounts of ζ-mRNA are also present throughout fetal life and ζ-globin is present in cord bloods of non-

thalassemic newborns. Throughout most of fetal life, HbF production predominates with a small amount of HbA. The different γ-globin chains are produced in a ratio of $^{G}\gamma$ to $^{A}\gamma$ of 3:1, which remains constant until late gestation. At mid-term, the bone marrow begins to take over as the major site of red-cell production, though erythropoiesis is also found in the spleen, as well as in other tissues. Towards the end of gestation there is a gradual and reciprocal switch from HbF to HbA production. At birth, cord blood normally contains ~70% HbF and this declines to ~ 20% by 3 months, 7.5% at 6 months, and less than 2% by the age of 1 year. At the same time there is the differential decline in $^{G}\gamma$- and $^{A}\gamma$-globin chain production so that their relative proportions become closer to equal. Both fetal and adult hemoglobins are produced in the same cell during the switching period, with a gradual increase in the proportion of cells containing predominantly HbA. The proportion of HbF continues to decline throughout childhood and probably throughout adult life, at which time the small amount of HbF is only detectable in 3-5% of red cells, known as F cells. The other hemoglobin produced in the adult red cells is HbA_2 ($\alpha_2\delta_2$), which forms 2-3% of total hemoglobin. Thus, in normal adults, three hemoglobins are generally found; HbA (97%), HbA_2 (2-3%) and HbF (<1%) (Figure 1.2 and table 1.1) (Weatherall, 1981).

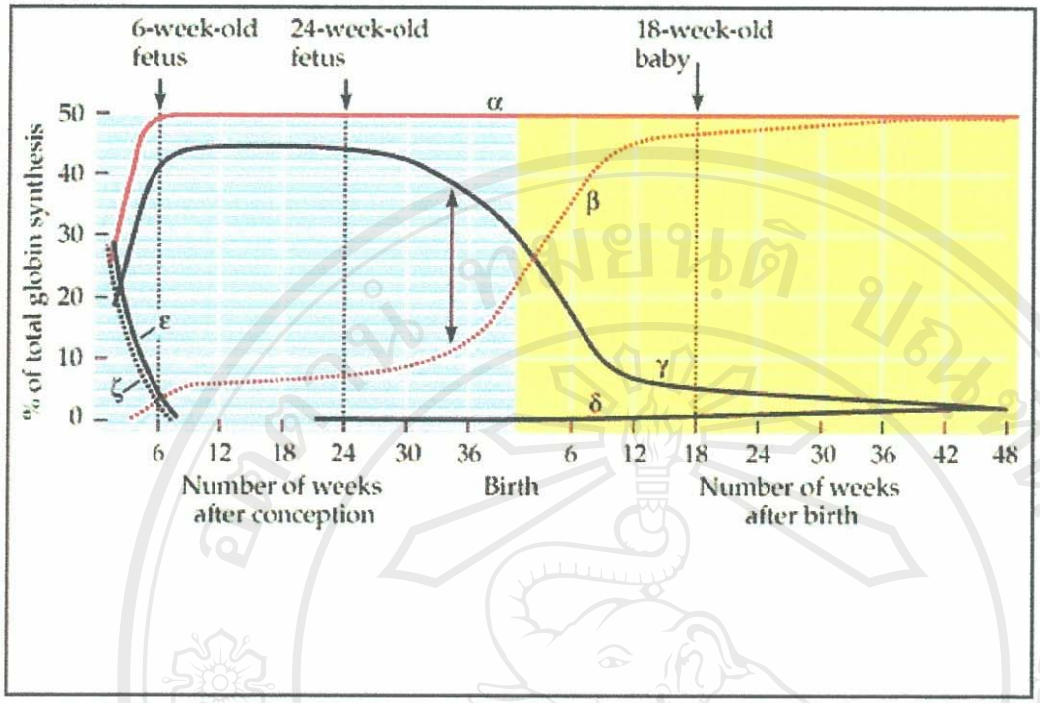

Figure 1.2 Normal developmental changes of globin chain synthesis and hemoglobin type in human.

Table 1.1 Hemoglobin type in human (Weatherall, 1981).

Hemoglobins	Embryo	Fetus	Adult
Gower 1 ($\zeta_2\varepsilon_2$)	42%	-	-
Gower 2 ($\alpha_2\varepsilon_2$)	24%	-	-
Portland ($\zeta_2\gamma_2$)	Present*	-	-
F ($\alpha_2\gamma_2$)	present*	90%	<1%
A ($\alpha_2\beta_2$)	-	10%	97%
A2 ($\alpha_2\delta_2$)	-	-	2.50%

* HbF and Portland have very similar mobility on cellulose acetate electrophoresis at pH 8.5, thus these two hemoglobins account for about 34% of total hemoglobins when blood samples from early embryo are examined by this technique.

1.4 The localization and organization of globin gene

The genes that regulate the synthesis and structures of the different globins are organized in two separate cluster (Smith, 1958). The α-like globin genes, which are encoded on chromosome 16, are found in the order 5'-ζ-ψζ-ψα-α_2-α_1-θ-3' (Lauer, *et al.,* 1980). The β-like globin genes, on chromosome 11, occur in the order 5'-ε-$^{G}\gamma$-$^{A}\gamma$-ψβ-δ-β-3' (Figure 1.3).

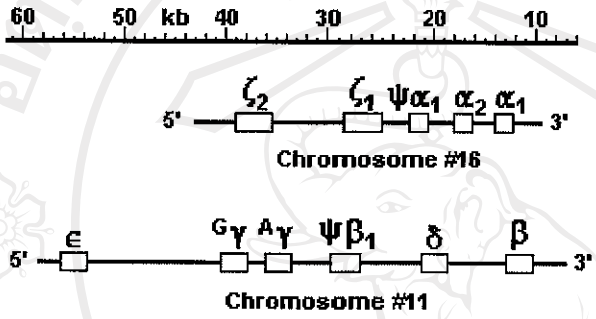

Figure 1.3 Localization and organization of α- and β-globin genes.
(http://www.people.virginia.edu/-rjh9u/gif/hgenemap.gif)

1.5 The α-globin gene cluster

The α-globin gene cluster occupies a region of about 30 kb close to the tip of the short arm of chromosome 16 (16p13.3). The α-globin gene family has evolved through a series of gene duplications and sequence divergences. The position of α-globin gene cluster arranged in the following order 5'-ζ-ψζ-ψα-α_2-α_1-θ-3'. The CAP site of ζ-globin gene is designated 0, the α_2 gene, at +20 kb, lies 20 kb away from the ζ gene, and the α_1 gene lies a furture 3.7 kb away, at +24 kb. The major α-globin gene regulatory element, HS-40, are centered on a region 40 kb upstream of the ζ-globin genes; the site of major regulatory element of the α-globin gene cluster. Structure of HS-40 is composed of 350 bp core fragment retaining most of the activity and

containing a duplicated NF-E2 binding site flanked by GATA-1 sites, as well as CACCC motifs and a YY1 binding site. The cluster contains three psudogenes ($\psi\zeta$, $\psi\alpha_2$, $\psi\alpha_1$) and another gene, θ. The latter is transcribed at a low level in erythroid cells (Figure 1.4) The α-globin genes are divided into two three exons by two non-coding sequences. The first exon encompasses 5' untranslated sequences (UTR) and the codons for amino acids 1-31 in the α-globin genes. Exon2 encodes 32-99, the third exon encodes the amino acids 100-141 (Figure 1.6) (Clegg, 1987).

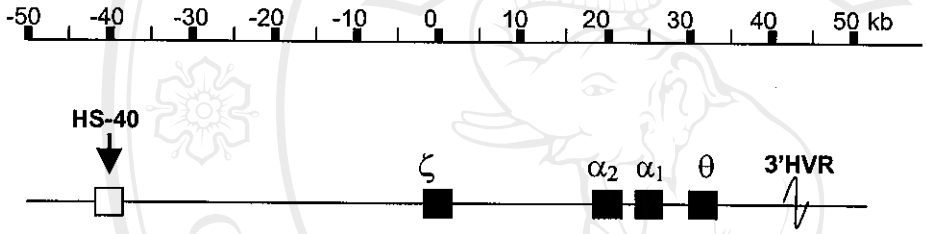

Figure 1.4 Schematic diagram of the human α-globin gene cluster in 16p13.3 (Weatherall, 2001).

1.6 β-globin gene cluster

The 60-kb β-globin gene cluster is localized on the short arm of chromosome 11 (11p15.5) containing the non α-globin genes arranged in the following order: 5'-ε-$^{G}\gamma$-$^{A}\gamma$-$\psi\beta$-δ-β-3'. There are two γ-globin genes, $^{A}\gamma$ and $^{G}\gamma$, which differ only in the codon 136 at which alanine is for $^{A}\gamma$ and glycine for $^{G}\gamma$ (Schroeder, *et al.*, 1971). The two fetal γ-globin genes lie 15 and 20-kb downstream from the embryonic ε gene, while the δ and β-genes are 35 and 43-kb further downstream. The β-pseudo gene ($\psi\beta$) is located between the $^{A}\gamma$-globin gene and the δ-globin gene. It is similar to the β-globin gene, but has been permanently altered by deletion and an internal stop codon, so that it cannot

code for a functional peptide. Upstream of the ε-globin gene lies the locus control region (LCR), the regulatory region essential for expression of all globin genes in the complex. It spans 15 kb and contains four elements (HS1 to HS4) which are marked by erythroid-cell-specific DNase1-hypersensitive sites (Tuan, 1985). There are two other hypersensitive sites, one 5' to the LCR and one 20 kb 3' to the β-globin gene domain (Figure 1.5). The structural sequence of β-globin genes included three exons separated by two introns, or interventing sequences (IVS) (Figure 1.6) (Lawn, *et al.*, 1980). From the CAP site, the start of transcription, the first exon encompasses 50 bp of 5' untranslated sequence (UTR) and the codons for amino acids 1-30 in the β-globin genes. Exon 2 encodes 31-104, the third exon encodes the amino acids 105-146. A promoter region located 5' to the encoding portion of the gene contains two and possibly more of the sequences. This region is believed to be the recognition and binding site for RNA polymerase which catalytes mRNA synthesis. The first of this promoter sequence, the ATA box or Goldberg Hogness box is located 20-30 bp 5' to the transcribed portion of the genes. The function of this sequence appear to be to locate and direct the precise site of initiation of transcription of the gene (Malthis, 1981). A second conserved sequence, the CCAAT box, is located 70-80 bp 5' to the transcription initiation site. A highly conserved hexanucleotide sequence AATAAA has also been identified in the 3' noncoding region of the globin genes, 15-19 bp 5' from the point of transcription termination (Michelson, 1980). The function of this sequence is as a termination signal for RNA polymerase, or as a recognition site for enzymes involved in the polyadenylation of mRNA (Proudfoot, 1976).

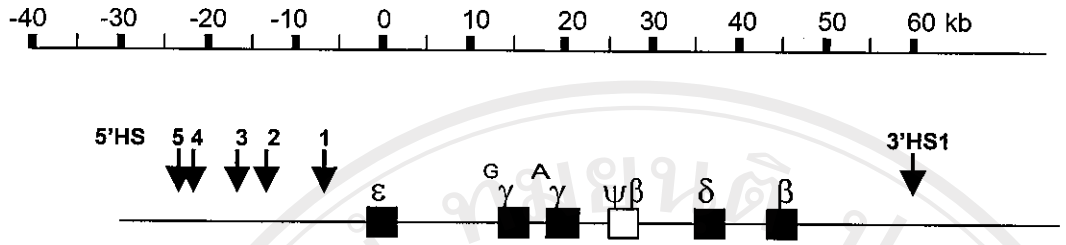

Figure 1.5 Schematic diagram of the human β-globin gene cluster in 11p 15.5 (Weatherall, 2001).

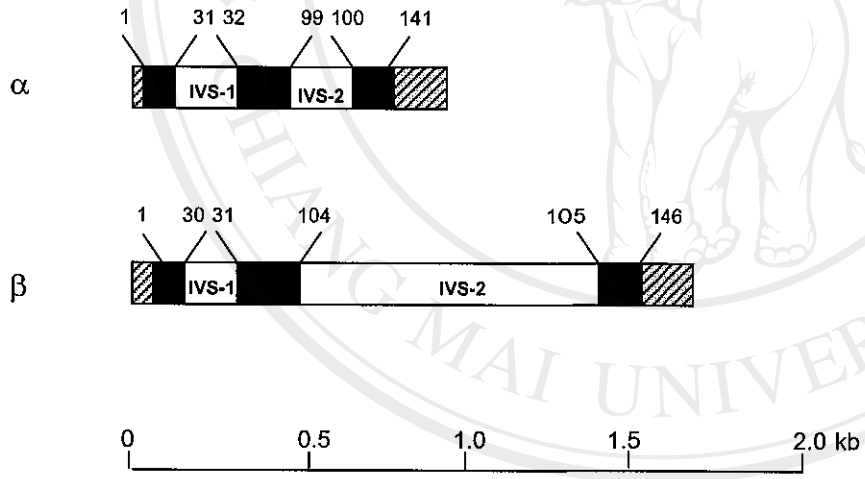

Figure 1.6 Structure of the human α- and β-globin genes. The black blocks represent the coding regions of the gene; the open blocks represent the intervening sequences (introns) IVS-1 and IVS-2. The numbers above the diagram indicate the amino acid codon positions of the coding sequence (Weatherall, 1981).

1.7 Thalassemia and hemoglobinopathies

The thalassemia is a group of disorders in which syntheses of globin chains are reduced or totally absent. The reduced or absent rate of synthesis of one or more of the globin chains leads to imbalanced globin-chain synthesis, defective hemoglobin production, and damage to the red cells or their precursors from the effects of the globin subunits that are produced in relative excess. This causes expansion of the ineffective marrow with severe effects on development, bone formation and growth. The major cause of morbidity and mortality is the effect of iron deposition in the endocrine organs, liver, and heart, which results from increased intestinal absorption and frequent of blood transfusions (Weatherall, 2001).

Abnormal hemoglobin, hemoglobinopathies, occurs from genetic alterations including point mutations, deletions or insertion of the globin genes. Over 90% of known variants have arisen by substitution of one amino acid residue in one chain type and over 60% involve the β-globin chain. Because an individual inherits only two β-globin genes, a β chain variant usually constitute about half of the total hemoglobin in the red cell. In contrast, most individual have four α-globin genes. Therefore, α-chain variants usually contribute only about 25% of the total hemoglobin and the consequence is generally milder than those produced by β-chain variants. Hemoglobin variants may cause clinical symptom or not depending on the site where the mutation occurs.

1.7.1 Common types of thalassemia and hemoglobinopathies

Two types of thalassemia commonly found in Thailand are alpha (α) and beta (β) thalassemia.

1.7.1.1 α-thalassemia

The α-thalassemia is characterized by a reduced or absent rate of α-globin chain production. DNA mapping has reveals that the gene deletion is the major cause of α-

thalassemia and that non-deletion is very rare. In α-thalassemia 1 (α^0-thalassemia) there is a deletion of both α_1 and α_2-globin genes from the α-globin gene cluster. The α-thalassemia 1 Southeast Asian (SEA) type is the one most frequently found in Thailand (Fischel-Ghodsian, *et al.*, 1988). The SEA deletion removes about 19.304 kb of α-globin gene cluster including α_1 and α_2-globin genes (Figure 1.7). In α-thalassemia 2 (α^+-thalassemia), one α-globin gene is functional. There are two types of α-thalassemia 2, one involving a deletion of 4.2 kb of DNA including the α_2–globin gene (leftward or 4.2-kb type) (Figure 1.8) (Embury, 1980; Trent, 1981) and another involving 3.7 kb of DNA between the duplicate α genes (rightward or 3.7-kb type) (Embury, 1980). The latter is more common in 10 east Asia and the point the of crossover which gives rise to it is located within the highly conserved region around the third exon of the α-globin genes, in which the homology is more than 99%. α-thalassemia is present throughout Southeast Asia. Its distribution is heterogeneous. In Thailand, the overall frequency of α thalassemia is 20-30%. The frequency of α-.thalassemia 1 is higher in the North than in the South; 10% in Chiang Mai and 3.5% in Bangkok whereas α-thalassemia 2 is between 16-20% (Fucharoen, 2002).

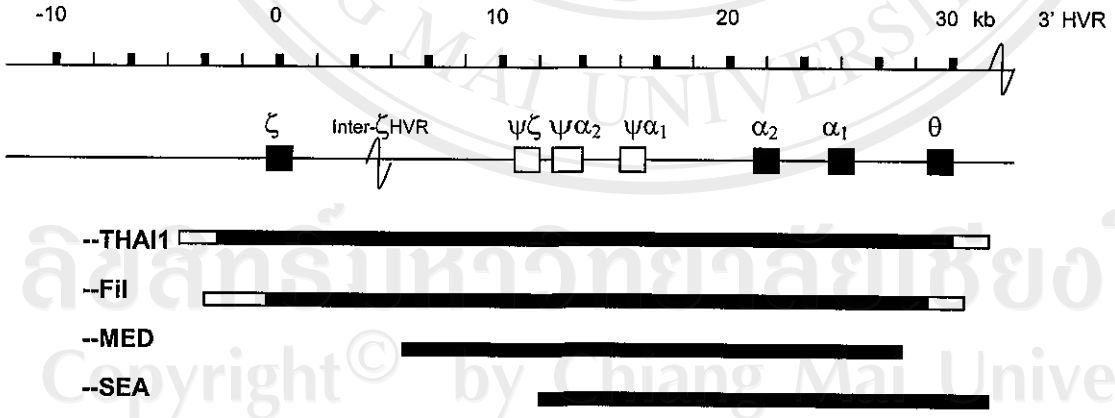

Figure 1.7 Deletions that cause α^0-thalassemia. Above : the α-gene complex is shown. Below: the extent of each deletion is shown by block bar (Fischel-Ghodsian, *et al.*, 1988; Pressley, *et al.*, 1980).

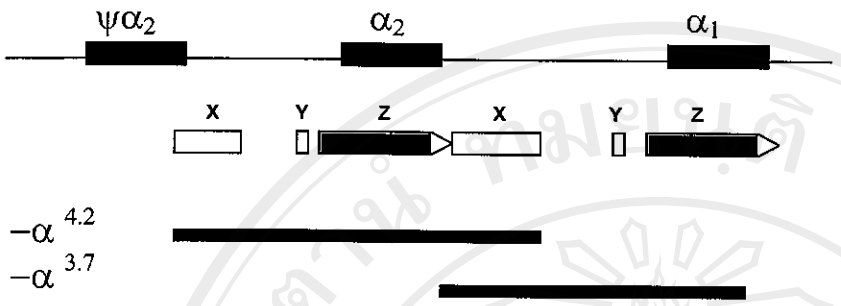

Figure 1.8 The deletions that underline the α^{+}-thalassemia. Above : the α-globin genes are shown with the duplication units divided into X, Y and Z boxes with regions of non-homology. Below: the extent of each deletion, represent by a black bar (Embury, 1980).

α-hemoglobinopathies

α-Hemoglobinopathy which is common in Thailand is hemoglobin Constant Spring (HbCS). HbCS is an α-globin chain variant with an elongated α-globin chain of 31 amino acids, which is resulted from point mutation of the termination codon of α_2-globin gene. Due to the instability of this elongated α-globin chains and the reduction of the expression of downstream α_1-globin, the clinical phenotype of those having HbCS allele is generally α-thalassemia 2. The percentage of HbCS is usually about 1% in heterozygous carriers, 5 to 7% in homozygotes and 3 to 5% in hemoglobin H disease from HbCS in-*trans* to deletional α-thalassemia. Heterozygotes of HbCS and two normal *trans* α-globin genes are hematologically normal. Homozygotes for HbCS have a mild hemolytic anemia and may have splenomegaly. Individuals with HbH disease from HbCS and two deletional α-thalassemia have a more severe disease with more HbH and Bart's than three-gene-deletional α-thalassemia (Figure 1.9). HbCS is

common in Southeast Asian and found in high frequency in American immigrants from some of those areas (Eckman, 1998).

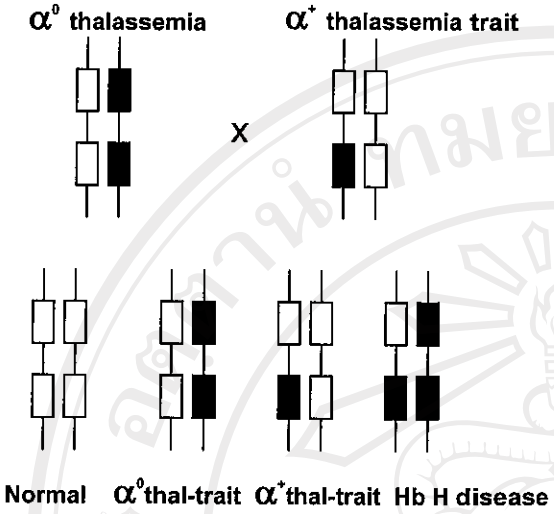

Figure 1.9 A representation of interactions of α^0 and α^+-thalassemia. The unshaded boxes represent normal α-globin genes, the dark-shaded boxes represent deleted or otherwise inactivated α-globin genes (Weatherall, 1981).

1.7.1.2 β-thalassemia

The β-thalassemia is a heterogeneous group of disorders characterized by decreased or absent β-globin chain synthesis. Point mutations and small deletions or insertions in the nucleotide sequences are mainly responsible for the molecular defects of β-thalassemia around the world and in Thailand. Two types of β-thalassemia are known; β^+-thalassemia, in which the production of the β chains is reduced, and β^0-thalassemia, in which the production of β-chain is totally eliminated. β^+-thalassemia generally involves defects at RNA processing or promoter region of the gene. In some cases it results from a mutation within the introns of β-globin gene. In β^0-thalassemia the absent of β chain synthesis is resulted from several causes such as a complete blockage at transcription or RNA processing, leading to lacking of the β globin mRNA

production. In some cases it is caused by point mutations in the DNA sequence, for instance, a nonsense mutation that provides a production of an incomplete β-globin chain. The frequency of β-thalassemia, in Thailand, varies from 1-9% (Fucharoen, 2002).

β-hemoglobinopathies

Hemoglobin E is the most common hemoglobin variant in the world population. In Thailand, it accounts for approximately 13% to 70% of the population (Chernoff, 1954; Sturgeon, 1955). G-A substitution at codon 26 of β-globin chain bring about the β-globin structural variant that when forms tetrameric structure with α-globin chain result in HbE. The β-globin chain codon 26 Glu-Lys mutation (GAG-AAG) partially activates a cryptic splice site towards the 3' end of exon 1, resulting in a proportion of abnormally splice mRNA (Orkin, *et al.*, 1982). Thus, less β^{E}-globin chain synthesized and mild thalassemia phenotype results (Traeger, *et al.*, 1980). The inheritance of HbE produces the phenotype of a mild form or β^{+}-thalassemia (Fucharoen, 2000).

1.8 The molecular defects of β-thalassemia

While gene deletion is the main cause of α-thalassemia, the molecular background of β-thalassemia is different. The majority of the β-thalassemia patients bears intact β-globin gene with minor change in the nucleotide sequence such as points mutations and small deletion of part of the gene.

1.8.1 Point mutations

Point mutation is the major cause of β-thalassemia, which results in either absence of the synthesis of β-globin chains; β^{0}-thalassemia, or reduction of synthesis; β^{+}-thalassemia. The majority of point mutations that affect β-globin gene function have been classified into 3 groups; transcriptional, RNA processing and RNA translational

mutations. ***First***, transcriptional mutations, this mutations involve single-base substitutions in the conserved DNA sequences in the β-globin promoter. Transient expression systems showed that the defects allow an output of β-globin mRNA ranging from 10% to 25% of the normal contributing to the relatively mild phenotype of β^+-thalassemia (Treisman, 1983). The transcription mutants emphasize the role of the conserved promoter sequences in the binding of transcription factors involved in the expression of the β-globin gene. Recently, a study has shown that the β-mutation at position –87 specifically ablates EKLF binding site resulting in the less efficient transcription of β-mRNA (Gregory, 1996). ***Second***, RNA processing mutations. The sequences critical for the splicing process include the dinucleotides GT at the 5' (donor) and AG at the 3' (acceptor) splice junctions between the exon and introns (Breathnach, 1981). Flanking these invariant dinucleotides are sequences that are fairly well conserved and a consensus sequence can be recognized at the exon-intron boundaries. They encompass the last three nucleotides of the exon and the first six nucleotides of the intron for the 5' donor site and the last 10 nucleotides of the intron and the first nucleotide of the exon for the 3' acceptor site. Both exons and introns also contain 'cryptic' splice sites which are sequences that mimic the consensus sequence for a splice site but are not normally used. During RNA processing the newly created site might be utilized preferentially leading to aberrant splicing. For example, mutations at position 5 of IVS1, G-T or A, considerably reduce splicing at mutated donor site compared with normal. The mutations appear to activate the use of three cryptic donor sites, two in exon1 and one in IVS1, which are utilized preferentially to the normal donor site (Treisman, 1983) and the substitution within IVS2: 654, C-T that generates new donor splice site (Cheng, *et al.*, 1984; Takihara, *et al.*, 1984). Substitutions or small deletions affecting the conserved AATAAA sequence in the 3' untranslated region result in ineffective clevage of the mRNA transcript and cause mild β^+-thalassemia (Antonarakis, *et al.*, 1984; Mantovani, *et al.*, 1988; Orkin, *et al.*, 1983; Poncz, *et al.*,

1982). ***Third***, mutants affecting RNA translation. Approximately half of the β-thalassemia alleles affect the different stages of RNA translation. In all this instances, no β-globin is produced, hence, causing β^{o}-thalassemia. Most of these defects result from the introduction of a premature termination codon owing to frameshifts or nonsense mutations. Frameshift defects; insertion or deletion of a single to a few nucleotides, alter the reading frame and cause premature termination further downstream. The examples of frameshif mutations are 4-bp (-TTCT) deletion at codons 41/42 and base addition (+A) at codons 71/72 (Kazazian, *et al.*, 1984; Sanguansermsri, *et al.*, 2001). In addition, the A-T substitution at codon 17 and G-T substitution at codon 43 generates stop codon that prematurely terminates β-globin mRNA translation (Tables 1.2, 1.3) (Fucharoen S, *et al.*, 1989). Globally, more than 600 β-globin gene mutations have been characterized todate, some of which are summarised in the table 1.2. In addition (http://globin.cse.psu.edu/cgi-bin/hbvar/counter), table 1.3 demonstrates the β-globin mutations identified among Thai β-thalassemic, in which defects of β-globin gene causing the disease in some of these patients remains to be determined.

Table 1.2 Summary of some molecular defects causing β-thalassemia (Weatherall, 1981).

Mutation	Type	Distribution
1. Transcriptional mutants		
Promoter regulatory elements		
-101 C-T	β^{++}	Mediterranean
-88 C-T	β^{+}	Asian Indian, US blacks
-86 C-A	β^{+}	Thai, Labanese
-28 A-G	β^{+}	Blacks, Southeast Asians
5' UTR		
CAP+1 A-C	β^{++}	Indian
CAP+40 to +43 (-AAAC)	β^{+}	Chinese
2. RNA processing		
Splice junction		
IVS1:1 G-T	β^{o}	Indian, Southeast Asians
IVS1:1 G-A	β^{o}	Mediterranean
IVS1:1 G-C	β^{o}	UAE
IVS2:1 G-A	β^{o}	Mediterranean, US blacks
IVS2:1 G-C	β^{o}	Iranian
IVS2: 849 A-G	β^{o}	US blacks
IVS2: 850 G-C	β^{o}	Yukoslavian
IVS2: 850 G-T	β^{o}	Japanese
IVS2: 850 –G	β^{o}	Italian
Consensus splice sites		
IVS1: 5 G-C	β^{+}	Indian, Southeast Asian
IVS1: 5 G-A	β^{+}	Mediterranean, Algerian
IVS1: 6 T-C	β^{++}	Mediterranean
IVS1: 128 T-G	β^{+}	Saudi Arabian
IVS2: 5 G-C	β^{+}	Chinese

Mutation	Type	Distribution
Cryptic splice sites in introns		
IVS1: 110 G-A	β^{+}	Mediterranean
IVS1: 116 T-G	β^{0}	Mediterranean
IVS2: 654 C-T	β^{0}/β^{+}	Chinese, Southeast Asians, Japanese
IVS2: 745 C-G	β^{+}	Mediterranean
Cryptic splice sites in exons		
CD10 GCC-GCA		Asian Indian
CD19 AAC-AGC	β^{++}	Southeast Asians
CD26 GAG-AAG	β^{+}	Southeast Asians, European
CD27 GCC-TCC	β^{+}	Mediterranean
3'-UTR		
RNA clevage: poly (A) signal		
AATAAA-AACAAA	β^{++}	US black
AATAAA-AATGAA	β^{++}	Mediterranean
AATAAA-AATAGA	β^{++}	Malay
AATAAA-AATAAG	β^{++}	Kurd
AATAAA-AA_ _AA	β^{+}	French, US Black
AATAAA-A_ _ _ _ _	β^{+}	Kurd
Others		
Terminal CD +6, C-G	β^{++}(silent)	Greek
Terminal CD +90, deletion 13 bp	β^{+}	Turkish
Terminal CD +47 C-G	β^{++}	Armenian
3. RNA translation		
Initiation codon		
ATG-GTG	β^{0}	Japanese
ATG-ACG	β^{0}	Yugoslavian

Mutation	Type	Distribution
ATG-AGG	β^0	Chinese
Nonsense codon		
CD7 GAG-TAG	β^0	English
CD15 TGG-TAG	β^0	Asian Indian, Japanese
CD15 TGG-TGA	β^0	Portuguese, Japanese
CD17 AAG-TAG	β^0	Chinese, Japanese
CD22 GAA-TAA	β^0	Reunion Island
CD26 GAG-TAG	β^0	Thai
CD35 TAC-TAA	β^0	Thai
CD37 TGG-TGA	β^0	Saudi Arabian
CD39 CAG-TAG	β^0	Mediterranean
CD43 GAG-TAG	β^0	Chinese, Thai
CD61 AAG-TAG	β^0	Black
CD90 GAG-TAG	β^0	Japanese
CD112 TGT-TGA	β^0	Slovenian
CD121 GAA-TAA	β^0	Czechoslovakian
Frameshift		
CD1-G	β^0	Mediterranean
CD2-4 (-9 bp, +31 bp)	β^0	Algerian
CD5 –CT	β^0	Mediterranean
CD6 –A	β^0	Mediterranean, US Black
CD8 –AA	β^0	Mediterranean, UK
CD8-9 +G	β^0	Asian Indian, Japanese
CD9-10 +C	β^0	Tukish
CD9-10 +T	β^0	Arab
CD14-15 +G	β^0	Chinese
CD15 –T	β^0	Asian Indian
CD16 –C	β^0	Asian Indian

Mutation	Type	Distribution
CD26 +T	β^0	Japanese
CD27-28 +C	β^0	Cninese, Thai
CD28-29 –C	β^0	Japanese, Egyptian
CD31 –C	β^0	Chinese
CD35 –C	β^0	Malay
CD36-37 –T	β^0	Kurdish, Iranian
CD37 –G	β^0	Kurdish
CD40-41 +T	β^0	Chinese
CD41 –C	β^0	Thai
CD42 –T	β^0	Japanese
CD41/42 –TTCT	β^0	Chinese, Southeast Asians, Asian Indian
CD42-43 +T	β^0	Japanese
CD42-43 +G	β^0	Japanese
CD44 -C	β^0	Kurdish
CD47-48 +ATCT	β^0	Asian Indian
CD71-72 +T	β^0	Chinese
CD71-72 +A	β^0	Chinese
CD72-73-AGTGA, +T	β^0	British
CD74-75 -C	β^0	Turkish
CD76 -C	β^0	Italian
CD84-85 +C	β^0	Japanese
CD88 +T	β^0	Asians Indian
CD95 +A	β^0	Southeast Asians

1.8.2 Deletions

In contrast to α-thalassemia, the β-thalassemia is rarely caused by major structural gene deletions. The 619-bp deletion (Orkin, 1980) removing the 3' end of the β-globin gene but leaves the 5' end intact was found in some Indian and Pakistanian living in Thailand in which it accounts for 20% of the β-thalassemia alleles (Thein, *et al.*, 1984). Recently, a 105-bp deletion of β-globin gene was characterized in Southern Thailand (Nopparatana, *et al.*, 1995). Moreover a 3.4-kb deletion, which removes the entrie β-globin gene was also found in Thailand (Table 1.2) (Sanguansermsri, *et al.*, 1990).

1.9 Clinical classification of β-thalassemia

The β-thalassemia can be classified on the clinical ground. The principle for this classification dose not rely solely on the molecular defects. Rather, several clinical and hematological aspects are considered before categorization of the patients. Using all the possible criterias, the β-thalassemia can be clinically divided into 3 groups including β-thalassemia major, β-thalassemia intermedia and β-thalassemia minor (Weatherall, 2001b).

1.9.1 β-thalassemia major is the most severe form of β-thalassemia. Genotypically, they can be either homozygous β-thalassemia or compound heterozygote for β-thalassemia mutations including HbE. The onset of the clinical symptom is usually before 2 years old. They show early growth retardation, pallor, icteric, and, as they grow older, brown pigmentation of the skin. Progressive expansion of the bone marrow in response to anemia lead to the classical thalassemia facies. A steady-state hemoglobin level is less than 6 g/dl (Weatherall, 1981). Their subsequent clinical courses and physical findings depend on adequate transfusion regimen or whether they are given

sufficient blood to develope normally. Without management, the patients of this type die prematurely at young age (Table 1.4).

1.9.2 β-thalassemia intermedia always present with a broad spectrum of clinical picture ranging from a condition little milder than from thalassemia major to symptomless disorder that is discovered only by chance examination of the blood (Ho, *et al.*, 1998). The age of presentation of β-thalassemia intermedia tends to be latter than β-thalassemia major (Modell, 1984). The clinical features are characterized by varying degree of anemia and splenomegaly and, in the more severe forms, bone changes similar to the major form of the illness. The hemoglobin level settles down to 7-10 g/dl (Ho, *et al.*, 1998). Blood transfution is not need to be as frequent as for those with thalassemia major. Intensive studies of the molecular pathology of this condition have provided some guidelines about genotype-phenotype relationships that are useful for genetic counselling (Table 1.4).

1.9.3 β-thalassemia minor are observed in these heterozygous for β-thalassemia. Usually they are characterized by mild anemia and small, poorly hemoglobinized red cells. Hemoglobin value of patients with β-thalassemia trait are usually in the range of 8-10 g/dl (Weatherall, 1981). The abnormallity is discovered only on performing a routine blood examination. Megaloblastic transformation due to folic acid deficiency occurs occasionally, particularly during pregnancy. There is mild degree of ineffective erythropoiesis (Table 1.4).

Table 1.3 β-thalassemia mutations in Thailand (Laig, *et al.*, 1989)

Mutations	Central (%)	North (%)	Northeast (%)	South (%)
β^0-thalassemia				
CD41/42 (-TTCT)	41.6	39.8	45.3	30.1
CD17 (A-T)	16.5	39.8	25.6	11.3
CD35 (C-A)	1.9			
IVS1 nt 1 (G-T)	1.3		3.4	6.0
CD71/72 (+A)	2.1		10.2	
CD8/9 (+G)				0.4
CD14/15 (+G)	0.3			
CD15 (-T)	0.3			0.4
CD26 (G-T)			0.9	
CD27/28 (+C)	0.5			
CD41 (-C)	0.5			1.4
CD43 (G-T)	0.5			
CD95 (+A)	0.3			
3.4-kb deletion	0.3	0.9		4.3
619-bp deletion	1.3			
1.5-bp deletion				0.4
CD123-125 (-ACCCCACC)			0.9	
IVS1 nt 1 (G-A)				0.4
β^+-thalassemia				
NT-86 (C-G)	0.3			
ATA NT-28 (A-G)	9.3	3.5	1.7	5.7
CD19 (A-G)	2.9			15.2
IVS1 nt 5 (G-C)	4.3	1.8	0.9	18.8
IVS2 nt 654 (C-T)	7.5	0.9	8.5	2.1
CD 26 (HbE)	0.5			
Unknown	7.8	13.2	2.6	3.1

Table 1.4 Summary of the criterias conventionally used for β-thalassemia clinical classification (Ho, *et al.*, 1998).

Criterions	β-thalassemia major	β-thalassemia intermedia
1. Clinical	severe	mild to severe
2. Hemoglobin	3-4 g/dl	7-10 g/dl
3. Frequency of blood transfusion	2-3 / months	occasional
4. Hepatomegaly	present	present (rare)
5. Splenomegaly	present	present

1.10 Characterized amoeliorating genetic factors on β-thalassemia

Recent progress towards a better understanding of the factors which can modify the phenotype of β-thalassemia has been intensively reviewed (Weatherall, 2001b). Three major alleviating genetic factors were revealed including the nature of β-thalassemia mutations, co-existance of α-thalassemia and co-inheritance of loci responsible for reactivation of HbF production (Camaschella, *et al.*, 1995; Rund, *et al.*, 1997). There are several particularly mild mutations such as IVS1 nt 6 (T-C) and those involving the regions -88 (C-T) and -28 (A-G) that are associated with mild clinical phenotypes (Weatherall, 2001). On the other hand, those having severe mutations such as 4-bp deletion at codons 41/42 or nonsense mutation at codon 17 are always accompanied by severe clinical phenotype. Co-inheritance of α-globin gene deletions such as -α/-α or --/αα or non deletional α_2-globin gene mutation in homozygous β^0-thalassemia is more likely to produce the clinical phenotype of thalassemia intermedia, whereas the co-inheritance of a single α-globin gene deletion in the same group of patients is usually associated with thalassemia major phenotype. In homozygous or compound heterozygous for β^+-thalassemia, co-inheritance of single α-globin gene

deletion is sufficient to produced and ameliorating effect (Fuebetta, *et al.*, 1983; Galanello, *et al.*, 1989; Thein, *et al.*, 1988; Wainscoat, *et al.*, 1983).

Although high HbF and γ-globin chains levels are always associated with mild clinical course of β-thalassemia; little is known about genetic variability in the production of hemoglobin F in this disorder. The most intensively and extensively analyzed genetic element that is linked to HbF augmentation is the C-T substitution at nucleotide position -158 in the promoter region of the $^{G}\gamma$-globin gene and creates an *Xmn*I cleavage site, thus, namely *Xmn*I$^{G}\gamma$ polymorphism (site) (Gilman, 1984; Labie, *et al.*, 1985). This polymorphism had been previously shown to be linked to increased expression of the $^{G}\gamma$-globin gene in β-thalassemia major and sickle cell anemia of diverse origins. The homozygous or heterozygous *Xmn*I$^{G}\gamma$ polymorphism might play an important role in underlying the milder forms of homozygous β-thalassemia in Afro-Asian populations (Kutlar, *et al.*, 1990). Later studies in Sardinian gave essentially similar finding. In the study of Ho and colleagues among the β-thalassemia of Asian-Indian background, it was clear that there was a strong association of the *Xmn*I-$^{G}\gamma$ polymorphism and high HbF level and mild clinical picture among β^{0}-thalassemia bearing IVS-1 (G-T) substitution (Ho, *et al.*, 1998). However, several studies has clearly shown that the *Xmn*I-$^{G}\gamma$ polymorphism is not the major *cis*-genetic determinant for HbF augmentation and that other putative genetic factors both in-*cis* or in-*trans* to the β-globin cluster may be invloved.

1.11 Background of the study

The β-thalassemia intermedia are extremely heterogenous, with respect to both their underlying genotypes and the spectrum of their clinical phenotypes, which range from a relatively severe degree of anemia requiring intermittent blood transfusion to an asymptomatic condition that is identified only by chance hematological study (Ho, *et al.*, 1998). The milder clinical expression of thalassemia intermedia is related to a lesser

imbalance of globin chain synthesis as compared to thalassemia major patients. This may result from inheritance of a mild β-thalassemia mutation, co-existance of α-thalassemia and the presence of loci responsible for HbF augmentation. Extensive data collected from the populations of the Mediteranian and Southeast Asia indicates that the co-existence of α-thalassemia may modify the phenotype of homozygotes or compound heterozygotes for different β-thalassemia mutation (Weatherall, 2001). For example, it is clear that the co-existence of the heterozygous stage for α-thalassemia 2 and the homozygous stage of β^0-thalassemia has little effect on the phenotype. On the other hand, individuals who are either α-thalassemia 1 heterozygotes or α-thalassemia 2 homozygotes, and who also are homozygous for β^+ thalassemia, may have a mild form of β-thalassemia intermedia. The genetic interaction which results in increased synthesis of fetal hemoglobin (HbF) also modifies the phenotype of β-thalassemia. The most intensively studied one is the upstream *Xmn*I-$^G\gamma$ polymorphism (Gilman, 1984). Although by itself, the *Xmn*I-$^G\gamma$ polymorphism dose not show significant functional relevance, but in the presence of hemopoietic stress it is always linked to elevated HbF production and mild clinical course.

Many surveys have clearly shown that different amoeliorating factors predominate in different population groups all over the world. In the study performed by Thein *et al.* (Thein, *et al.*, 1988) in Asian Indian β-thalassemia living in the UK, β-mutations were found to be the major modifying factor. In contrast, β-thalassemia mutations and α-globin deletion were shown to be the major amoeliorating factors among Chinese, Korean and Italian β-thalassemia (Camaschella, *et al.*, 1995; Chen, *et al.*, 1997; Chen, *et al.*, 2000). Moreover, the surveys performs by Fonseca *et al.* and Qatanani *et al.* (Fonseca, *et al.*, 1998; Qatanani, *et al.*, 2000) have demonstrated the close association of the mild β-thalassemia mutations as well as *Xmn*I-$^G\gamma$ polymorphism and mild clinical picture of β-thalassemia.

On the other hand, the studies conducted by Galanello *et al.* (Galanello, *et al.*, 1989) among Sardinian β-thalassemia failed to demonstrate any major alleviating factors; i.e. they act equally. The similar findings were also found in the survey among African American, Italian and Thai carried out by Ballas, Leoni and Winichagoon, respectively (Ballas, *et al.*, 1997; Leoni, *et al.*, 1991; Winichagoon, *et al.*, 2000).

Evidence of intra-country heterogeniety of these alleviating factors was revealed in the study in Indians subcontinent. In the survey among β-thalassemia in Panjab and Maharash states, Garewel and Colleagues found that mild β-thalassemia mutations and co-existence of α-thalassemia were closely associated with β-thalassemia intermedia (Garewal, *et al.*, 1994). However, surveys in other parts of country by Nadkarni *et al.* (Nadkarni *et al.*, 2001) showed that most of β-thalassemia intermedia results from co-existence of α-thalassemia and *Xmn*I-$^{G}\gamma$ polymorphism.

In Thailand, the incidence of β-thalassemia is 3-9% (Fucharoen, 2002) and that of HbE, the commonest β-structural variant in Thailand is 13-70% (Fucharoen, 2002). Homozygotes for β-thalassemia alleles and compound heterozygotes for β-thalassemia and HbE alleles result in β-thalassemia disease. In the study conducted by Winichagoon *et al.* (Winichagoon, *et al.*, 2000) among Thai homozygous β-thalassemia patients and by Fucharoen *et al.* (Fucharoen, *et al.*, 2000) among Thai HbE/β-thalassemia in the central Thailand, two type of the patient were clinically classified, including β-thalassemia major and β-thalassemia intermedia. These studies showed that all these amoeliorating factors act equally in modifying clinical outcome of the studied patients.

At Maharaj Nakorn Chiang Mai Hospital, patients with homozygous β-thalassemia and HbE/β-thalassemia attending the Pediatric Thalassemia Clinic are of the Northern Thai origin. Two kinds of patients; i.e. β-thalassemia major and β-thalassemia intermedia are encountered. However, the molecular background associated with the β-thalassemia intermedia and the β-thalassemia major has never been identified in this

center. Without genetic information responsible for the clinical outcome of the β-thalassemia intermedia, genetic counselling and prenatal diagnosis of those at risk for homozygous β-thalassemia and HbE/β-thalassemia are difficult to perform. Thus, characterization of genetic background of the β-thalassemia and identification of the genetic modifiers associated with the β-thalassemia intermedia in these patients would be extremely advantageous.

1.12 Objectives

1. To characterize molecular background of β-thalassemia intermedia and β-thalassemia major in Maharaj Nakorn Chiang Mai Hospital.
2. To evaluate the effect of genetic modifying factors on the clinical severity of β-thalassemia at Maharaj Nakorn Chiang Mai Hospital.