CHAPTER 1

INTRODUCTION

1.1 Statement and significance of the problems

Alpha thalassemia is caused by deficient synthesis of globin chains. A wide variety of point mutations and deletions has been described which decrease α - globin gene expression in humans. Clinical phenotype of the carriers varies according to the number of affected genes. Alpha thalassemia 2, the carriers of three α - globin genes (-α/αα) present with no detectable red blood cell abnormalities or globin chain imbalance, while α - thalassemia 1, the carriers of two function α - globin (-- / αα) or homozygous α- thalassemia 2 (-α / -α) have mild microcytic, hypochromic anemia with normal hemoglobin A 2 levels. The interaction of α - thalassemia 1 and α - thalassemia 2 result in hemoglobin H disease which is carries only one functional α - gene (- - / - α). Hemoglobin H disease is the most severe form of α - thalassemia phenotypes compatible with life, a condition characterized by anemia of variable severity, marked hypochromia and microcytosis. Inheritance of no functional α - globin gene (--/--) is usually incompatible with life an leads to Hb Bart's hydrops fetalis . Through out the world, α - thalassemia 2 occurs more frequently than any other type of thalassemia, it has a frequency of up to 30 % in certain parts of Africa, while in the other parts of the world, such as Polynesia and remote part of India, α - thalassemia 2 is inherited by more than 50 % of all individuals. The less frequently occurring nondeletional α- thal ($\alpha^T\alpha$ or $\alpha\alpha^T$) involves point mutations or other nucleotide alterations in the structural gene.

The thalassemia syndrome is the major health problem in Thailand. To approach this problem, population screening, genetic counseling and prenatal diagnosis programs were set to control the distribution of thalassemia. Large scale population screening programs to detect the carries of α - thalassemia 2 determinants have not been made in

many part of the world because the rather laborious and expensive DNA mapping was the only available technique.

Traditionally, identification of α - thalassemia 2 carriers may be achieved by globin chain synthesis analysis and by Southern blot analysis. Both procedures are relatively complex, time consuming, and cumbersome. Moreover, they imply the use of radioactive reagents and for these reasons are not practical and suitable for most laboratories. In this study, polymerase chain reaction (PCR) techniques were applied to detect the most common form of α - thalassemia 2 carriers, rightward type ($-\alpha^{3.7}$) in Northern Thai population. This method is simple, reliable and suitable for population screening and routine diagnosis.

มหาวิทยาลัยเชียงใหม่ Chiang Mai University

1.2 Literature review

1.2.1 Hemoglobin structure and function

Human hemoglobin is an approximately 64,000 Daltons compact globular protein (Adair,1925; Svedberg et al .,1926), and an elliptical shape with molecular dimension of approximately 64 x 55 x 50 A° (Pertz et al.,1963). The hemoglobin molecules are compose of four polypeptide chains and four heme prosthetic groups, in which the irons atoms are in the ferrous [Fe^{2+}] state. The protein portion, called globin, consists of two α- or α- like globin chains and two β- or β - like globin chains.

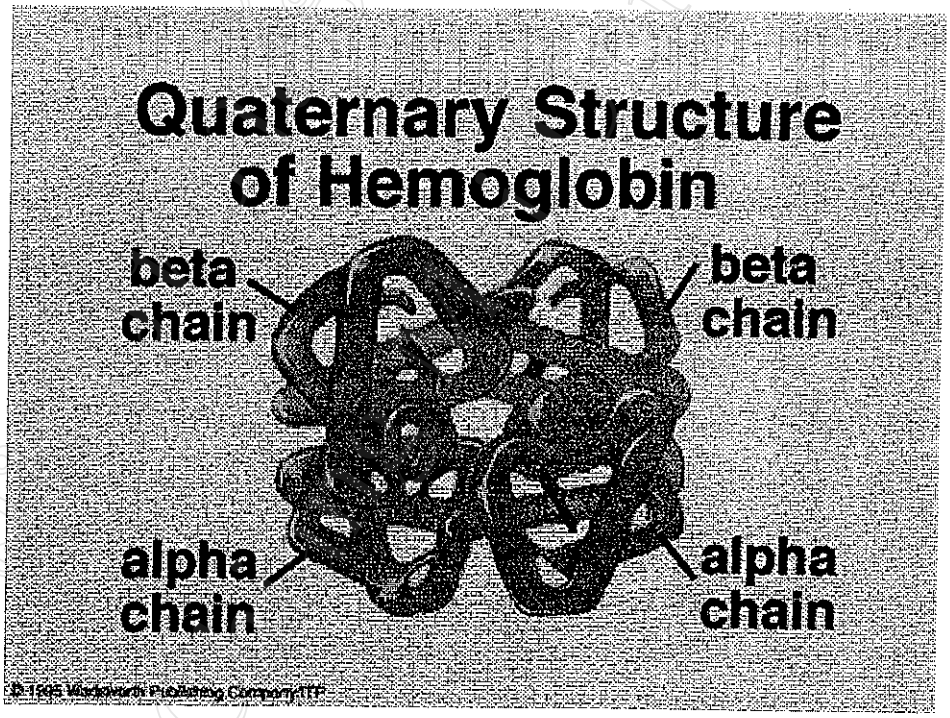

Figure 1. Hemoglobin structure

The hemoglobin is the most abundant blood protein in human, and represents more than 95% of the soluble protein content of the erythrocytes. The primary functional role of the hemoglobin is an oxygen carrier, it must be able to bind efficiently at the oxygen tension of the alveolar capillaries, and unload the oxygen to the tissues (Honig and Adams III, 1986). Allosteric interactions among the four globin chains result in the physiologically appropriate sigmoid oxygen binding curve and adjustments in the oxygen binding affinity to local tissue pH (Bohr effect), temperature, pCO_2 and levels of erythrocyte 2,3 diphosphoglycerate (Bunn and Forget, 1986). The function of the erythrocyte is therefore dependent upon balanced synthesis of α- and β-globin chains and their subsequence assembly into the functional hemoglobin tetramer (Liebhaber,1989).

1.2.2 Genetic control and synthesis of hemoglobin

Human adult hemoglobin is a heterogeneous mixture of proteins consisting of a major component, hemoglobin A, and a minor component, hemoglobin A_2 , consisting a bout 2.5 percent of the total. In intrauterine life, the main hemoglobin is hemoglobin F. The structure of these hemoglobins is similar. Each consists of two separate pairs of identical globin chains. Except for some of the embryonic hemoglobins, all the normal human hemoglobins have one pair of α chains : in hemoglobin A , these are combined with β chains ($\alpha_2\beta_2$), in hemoglobin A_2 with δ chains ($\alpha_2\delta_2$), and in hemoglobin F with γ-chains ($\alpha_2\gamma_2$).

Before the eight week of intrauterine life there are three embryonic hemoglobins, hemoglobins Gower1($\zeta_2\varepsilon_2$), Gower 2 ($\alpha_2\varepsilon_2$), and Portland ($\zeta_2\gamma_2$). The ζ and ε chains are the embryonic counterparts of the adult α and β and γ and δ chains, respectively. During fetal development there is an orderly switch from ζ to α and from ε to γ chain production, followed by β and δ chain production after birth.

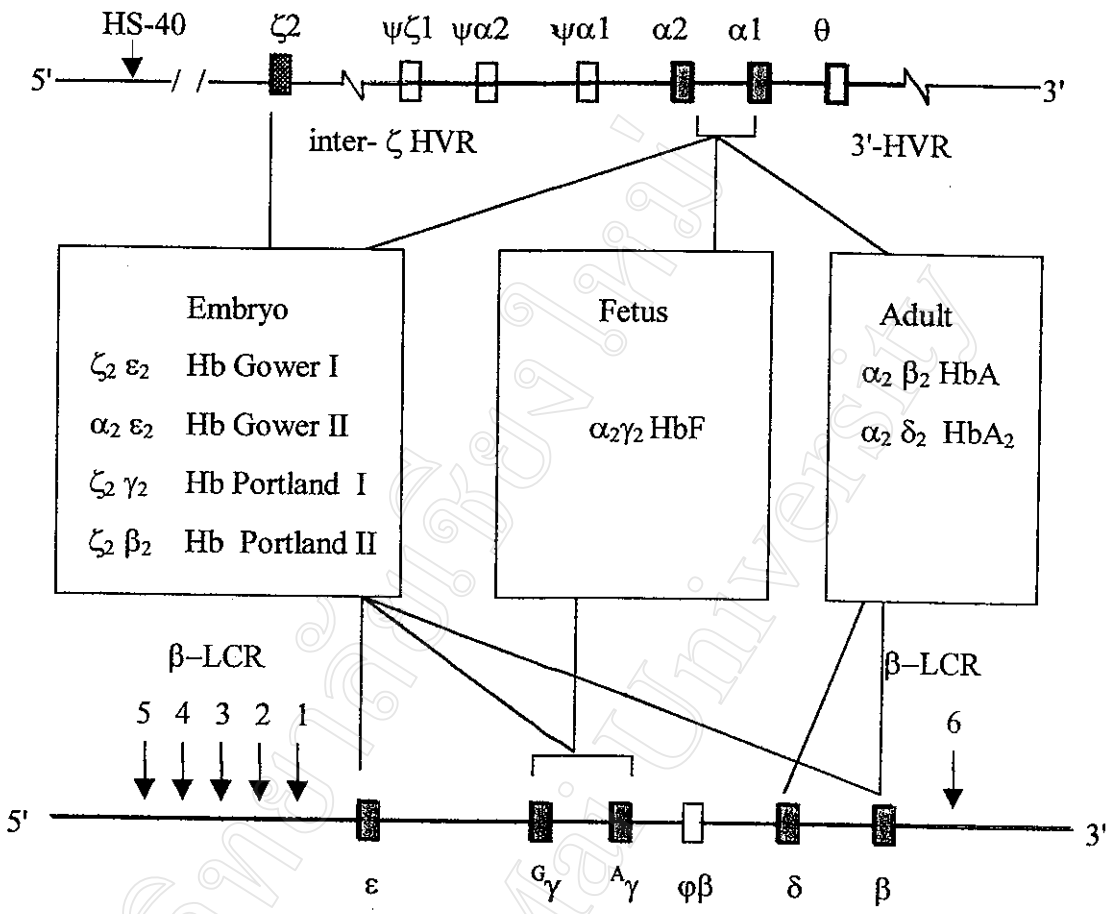

Figure 2. Schematic representation of the α globin (upper) and β globin (lower) gene clusters and their chromosomal location. The embryonic, fetal and adult hemoglobin code by the different genes during development is indicated in the flame between the clusters. Genes are shown as full boxes, psuedogenes as open boxes. The θ gene of undetermined function is indicated in gray. HVRs are indicated in zigzag lines. The position of the regulatory elements HS-40 and the β-LCR, consisting of hypersensitive sites 1-6, are indicated by vertical arrows (according to Bernini,1998).

1.2.3 Globin gene clusters

Alpha globin gene cluster (Weatherall,1995)

Although there is some individual variability, the α gene cluster usually contains one functional ζ gene and two α genes, designated α1 and α2. It also contains four psuedogenes ψζ1, ψα2, ψα1 and ψθ1 (Bunn and Forget, 1986, Weaterall, 1981). The latter has been discovered only recently and is remarkably conserved among different species. Although it appears to be expressed in early fetal life, its function is unknown; it seems likely that it can produce a viable globin chain. Each α gene is located in a region of hematogy approximately 4 kb long, interrupted by two small non homologous regions (Orkin, 1978., Lauer, 1980., Liebhaber, 1981.). It is thought that the homologous regions have result from gene duplication and that the nonhomologous segments may have arisen subsequently by insertion of DNA in to the noncoding regions around one of the two genes. The exons of the two α globin genes have identical sequences. The first intron in each gene is identical but the second intron of α1 is nine bases longer and differs by three bases from that in the α2 gene(Liebhaber,1980,Liebhaber,1981., Proundfoot, 1980). Despite their high degree of homology, the sequences of two α globin genes diverse in their 3' untranslated regions 13 bases beyond the TAA stop codon. It appears that the production of α2 messenger RNA exceeds that of α1 by a factor of 1.5 to 3. The ζ1 and ζ2 genes are also highly homologous. The introns are much larger than those of α - globin genes, in contrast to the latter, IVS - 1 is larger than IVS - 2. In each ζ gene, IVS-1 contain several copies of a simple repeated 14 - bp sequence that is similar to sequences located between the two ζ genes and near the human insulin gene. There are three bases changes in the coding sequence of the first exon of ζ1 , one of which gives rise to a premature stop codon, thus making it an inactive psuedogene. The regions separating and surrounding the α - like structural genes have been analyzed in detail. Of particular importance to the thalassemia filed is the fact that this gene cluster is highly polymorphic(Higgs et al.,1986). There are five hypervariable regions in the cluster, one downstream from the α1 gene, one between

the ζ and ψζ genes, one in the first intron of both the ζ genes, and one 5' to the cluster. These regions have been found to consist of varying numbers of tendem repeats of nucleotide sequences. Taken together with the single base restriction fragment length polymorphisms (RFLP), the variability of the α - gene cluster reaches a heterozygosity level of approximately 0.95. Thus it is possible to identify each parental α - globin gene cluster in the majority of persons. This heterogeneity has important implications for tracing the history of the thalassemia mutations .

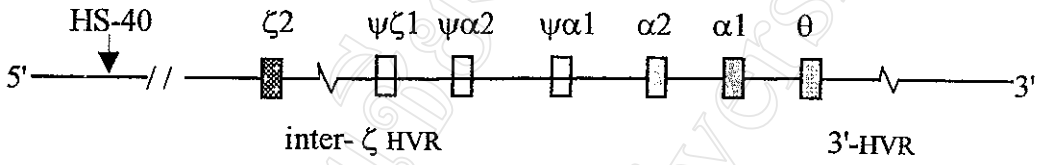

Figure 3. Schematic representation of the α - globin gene clusters (according to Bernini,1998)

Beta globin gene cluster (Weatherall, 1995)

The arrangement of the β - globin cluster on the short arm of chromosome 11 is shown in figure 4. Each of the individual genes and their flanking regions have been sequenced (Fritsch,1980., Spritz, 1980., Baralle, 1980., Slightom, 1980.). Like the α1 and α2 gene pairs, the $^{G}\gamma$ and $^{A}\gamma$ share a similar sequence. In fact, the $^{G}\gamma$ and $^{A}\gamma$ genes on one chromosome are identical in the region 5' to the center of the large intron yet show some divergence 3' to that position. At the boundary between the conserved and divergent regions there is a block of simple sequence that may be a hot spot for the initiation of recombination events that have led to unidirectional gene conversion.

Like the α - globin gene, the β-globin cluster contains a series of single point RFLPs, although in this case no hypervariable regions have been identified(Jeffrey, 1979., Antonarakis, 1982). The arrangement of RFLPs, or haplotypes, in the β-globin gene cluster falls into two domains. On the 5' side of the β gene, spanning about 32 kb from the ε gene to the 3' end of the ψβ gene, there are three common patterns of RFLPs. In the region encompassing about 18 kb to

the 3'side of the β-globin gene there are three common patterns in different populations. Between these region there is a sequence of about 11 kb in which there is randomization of the 5' and 3' domains and hence where a relatively higher frequency of recombination may occur. The β-golbin gene haplotypes are similar in most population but differ markly individuals of African origin; these results suggest that these haplotype arrangements were laid down very early during evolution, and they are consistent with data obtained from mitochondrial DNA polymorphisms that point to the early emergence of a relatively small population from Africa with subsequent divergence into other racial groups (Wainscoat, 1982.). Again. they are extremly useful for analyzing the population genetics and history of the thalassemia mutations. The regions flanking the coding regions of the globin genes contain a number conserved sequences that are essential for their expression(Nienhuis, 1987.). The first is the ATA box, witch serves accurately to locate the site of transcription initiation at the CAP site, usually about 30 bases downstream, and also appears to influence the rate of transcription. In addition, there are two so-called upstream promoter elements: 70 or 80 base pairs (bp) upstream is a second conserved sequence, the CCAAT box, and further 5', approximately 80 to 100 bp from the CAP site, is a GC-rich region with a sequence that can be either inverted or duplicated. These promoter sequences are also required for optimal transcription, and as we shall see later, mutations in this region of the β-globin gene cause its defective expression. The globin genes also have conserved sequences in their 3' flanking, notably AATAAA, which is the polyadenylation signal site.

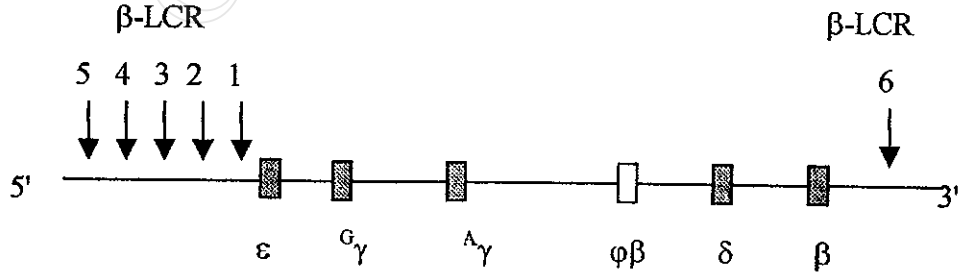

Figure 4. Schematic representation of the β - globin gene clusters (according to Bernini,1998)

1.2.4 The thalassemia

The thalassemias are the commonest single gene disorder. The name 'thalassemia' , which is derived from the Greek word for sea, reflects the easy notion that is a disorder of Mediterranean populations. This condition occurs widely throughout the Mediterranean region, Africa, the Middle East, the Indian subcontinent, and Southeast Asia. This family of genetic disorders, all characterized by a reduced rate of production of one or more the globin chain of hemoglobin, is divided in to α, β, $\delta\beta$, $\varepsilon\gamma\delta\beta$ thalassemias, depending on which globin chain is affected. The hallmark of thalassemia is imbalanced globin chain synthesis. The clinical consequences of hemoglobin and of the deleterious effects of the chains which are produced in excess on red cell maturation, survival and function [Weterall,1985].

1.2.5 The α-thalassemias and it molecular defects

The α-thalassemia is a hemolytic anemia resulting from deficient synthesis of α -globin. The deficiency of α-globin result insufficient production of functional hemoglobin and in the accumulation of unstable Hb Bart's or Hb H tetramers with consequent accelerated red cell destruction. α-thalassemia and can result from defect(s) in, or deletion(s) of, one or more of the four α-globin genes(Liebhaber, 1989). There are two main groups of α- thalassemia determinants. First, there are the α° thalassemia (formerly called α - thalassemia 1), in which no α - chains are produce from an affected chromosome. That is both like α- globin genes are inactivated. Second , there are the α^{+} thalassemias (formerly called α - thalassemia 2), in which the output of one of the linked pair of α- globin genes is defective. The α^{+} thalassemias are subdevided into deletion and nondeletion types. Both the α^{0} thalassemias and deletion and nondeletion forms of α^{+} thalassemia are all extremely heterogeneous at the molecular level (Weatherall, 1995).

1.2.5.1 α° thalassemia or α - thalassemia 1 due to total deletion of two α - globin genes per chromosome designated (--/)

The most common deletions, (--SEA) and (-- MED), occur in Southeast Asia and the Mediterranean Basin , respectively. These two deletions, remove both α - globin genes but spare the functional ζ gene (Nicholls et al., 1987). The deletion that causes the common form of α° thalassemia in southeast Asia (--SEA) extent upstream from near the 3' hypervariable region to the region close to the third exon of the $\varphi\zeta$ - gene. The deletion that produce the common Mediterranean from of α° thalassemia starts in approximately the same 3' region as in southeast Asian counterpart, but extends upstream to just 5' to *Bgl* II site in the $\varphi\zeta$ - gene. The extent of the other, less common, deletions that produce the α° thalassemia phenotype. With the exception of the $(\alpha\alpha)^{RA}$ mutant (a deletion from the α - cluster in which both α - genes remain intact), all of the deletions described in this section either completely or partially [- $(\alpha)^{5.2}$ and -$(\alpha)^{20.5}$] delete both α- globin genes, and therefore no α- chain synthesis is directed by these chromosomes in vivo. Those α- thalassemia defects that have been fully characterized result from illegitimate or nonhomologous recombination events (Nicholls et al., 1987). The deletions range in size from rather small 5.2 kb to those which remove the entire cluster [-- Fil, -- Thai (Winichagoon et al., 1984) and a new deletion of > 47 kb in a Northern European family (Fischel - Ghodsian et al ., 1988 ; Fortina et al., 1988)].

The two most common deletions, (--SEA) and (--MED), occur in Southeast Asia and the Mediterranean Basin, respectively. These two along with two deletional deletions (--SEA, and ---$(\alpha)^{20.5}$) are all approximately the same size (20 - 30 kb) and remove both α- globin genes (--) but leave the functional ζ2 gene(Nicholls et al., 1987).

1.2.5.2 α^+ thalassemia or α- thalassemia 2 : one deleted α - globin gene per chromosome designated (-α/).

There are two common deletion forms of α^+ thalassemia that are designated by the size of the deletion as $\alpha^{3.7}$ and $\alpha^{4.2}$ .The most common form is the $\alpha^{3.7}$ deletion results in the lost of 3.7 kb of DNA, while a less common form is the $\alpha^{4.2}$ deletion result in the lost of 4.2 kb of DNA. Both of these deletions results form unequal homologous recombination within the α- globin cluster. The two globin genes are imbedded in a large region of homolgy which is divided by short divergent regions into three homology subsegments ; X, Y and Z (Lauer et al., 1980). These regions of homology within and surrounding the two α-globin genes can mediate homologous unequal recombination. The cross over can occur between the two X-regions or the two Z- regions in the 4.2 kb and 3.7 kb deletions, respectively. Because of the relative position of these crossovers, $\alpha^{3.7}$ is often referred to as the rightward deletion and the $\alpha^{4.2}$ as the leftward deletion.

The $\alpha^{3.7}$ deletion can be further subdivided based on the exact position of crossing over within the Z box (Higgs et al., 1984). Type I occurs within the large region of homology extending form 863 nucleotide 5' of the two α- globin genes to the 7 base insertion / deletion divergence within the second intron. This is by far the most common form. Type II and III crossovers occur within the homology blocks extending between the second intron and the divergent segment in the 3'- non translated region of the third exon and in the small segment of homology surrounding the polyA addition site, respectively.

The $\alpha^{3.7}$ type I deletion is the extremely wide spread. The gene frequency of the $\alpha^{3.7}$ mutation can reach very high levels. In the Mediterranean basin the prevalence of the heterozygous state of $\alpha^{3.7}$ is approximately 5-10 % (Kanavakis , 1988), in southeast Asia approximately 10-20 % (Nicholls , 1987), in the certain region of West Africa the frequency approximates 20-30 % (Dozy et al., 1979), and in specific areas of India and Papua New Guinea the incidence is as high as 90% (Oppenheimer , 1988).

The $-\alpha^{4.2}$ is frequently found in Asian and related populations although it has also been reported in numerous other populations (Pacific Islannders, Black, Mediterranean, etc.). The linkages that of these mutations to specific α - globin gene cluster haplotypes suggests that each has occurred by a number of independent mutational events with subsequent positive selection in certain populations (Yenchitsomanus ,1985).

มหาวิทยาลัยเชียงใหม่ Chiang Mai University

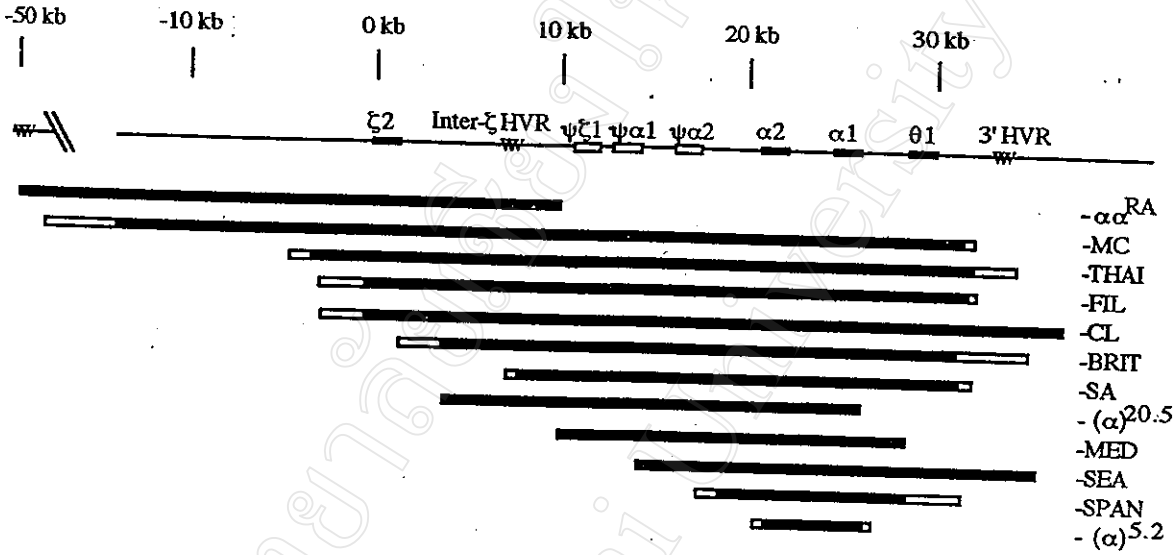

Figure 5. The deletion of the alpha globin gene cluster that are responsible for α° thalassemia

มหาวิทยาลัยเชียงใหม่ Chiang Mai University

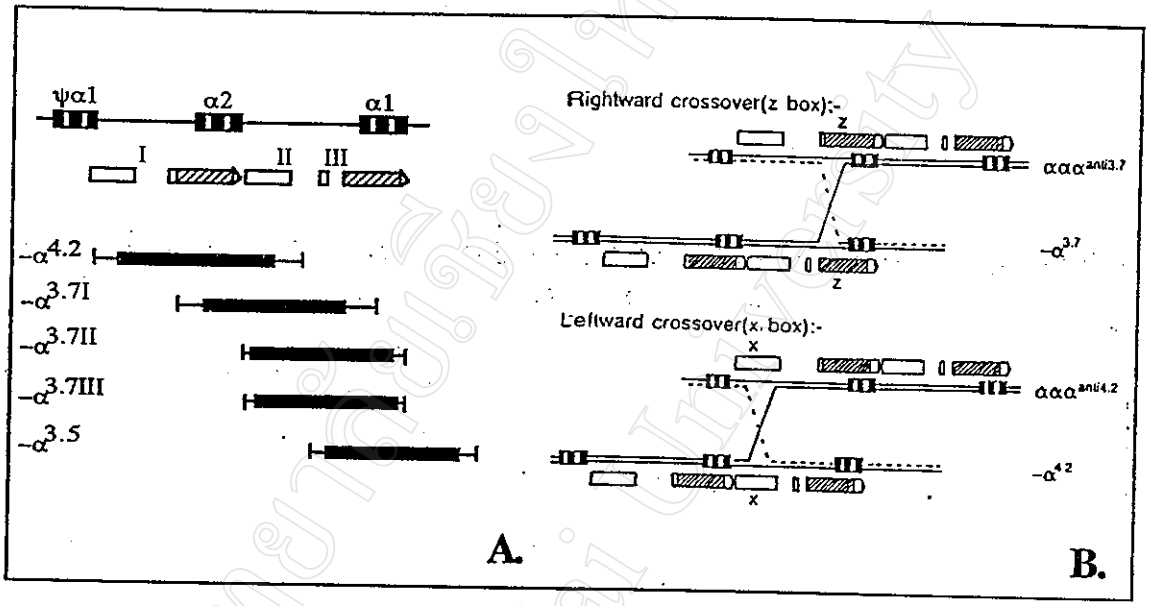

Figure 6. The single alpha - globin gene deletion and its proposed mechanisms (according to Higgs et al., 1989). A : The duplicated X, Y, and Z box arrangement containing the alpha-globin genes. Nonhomologous regions (I, II, and III) are indicated. The extend of each deletion is indicated by the solid boxes and the limits of the break points are represented by solid lines. B: The proposed crossovers between the homologous regions that lead to the $\alpha^{3.7}$ and $\alpha^{4.2}$.

1.2.5.3 Nondeletion α - thalassemia

Since expression of the α 2 gene is two to three times greater than that of the α 1 gene, it is not surprising that most of the nondeletion mutants discovered to date affect predominantly the expression of the α 2 gene; presumably this is ascertainment bias because of the greater phenotypic effect of these lesions. It is also possible that they have come under greater selective pressure.

Alpha thalassemia mutations can be classified according to the level of gene expression they affect. Two processing mutations have been identified. The first consists of a pentanucleotide deletion including the 5' splice site of IVS-I of the α2-globin gene. This involves the invariant GT donor splicing sequence and thus completely in activates the α2 gene. The second mutant of this type, found commonly in the Middle East (Thein, 1988), involves the poly A addition signal site (AATAAA→ AATAAG) and downregulates the α2 gene by interfering with 3' end processing.

A second group of nondeletion α- thalassemias resulted from mutations that interfere with the translation of messenger RNA. In one case the initiation codon is inactivated by a T→ C transition, and in another case efficiency of initiation is reduced by a dinucleotide deletion in the consensus sequence around the start signal. Four mutations that affect termination of translation and give rise to elongated α - chains have been identified: hemoglobins Constant Spring, Icaria, Koya Dora, and Seal Rock. Each specifically changes the termination codon TAA so that an amino acid is instead of the chain terminating. This is followed by read-through of messenger RNA that is not normally translated until another " in phase " stop codon is reached. Thus each of these variants has an elongated α - chain. It seems likely that the " read through " of α- globin messenger RNA that is usually not utilized somehow reduces it stability. There is one nonsense mutation, in exon 3 of the α2 globin gene. There are four structural mutations that cause α - thalassemia. In each case a highly unstable α - globin chain is produced; these are hemoglobins Quong Sze (Liebhaber,1983) , Suan Dok (Sanguansermsri,1979), Petah Tivah (Honing, 1981), and Evanston (Honing, 1984).

A number of mutation within the α-globin have been clearly etiologic in α - thalassemia. This non deletion category included both single base substitutions as very small deletions. This is summarized in table 1.

	Affected Gene	Affected Sequence	Mutation	Geographical Distribution	comments
RNA	α2	IVS I donor site	GAGGTGAGG->GAGG	Mediterranean	aberrant splicing
Processing	α2	poly(A) signal	AATAA->AATAAG	Mediterranean, Middle East	↓ efficiency of 3'end processing
	α2	Initiation codon	ACCATGG->ACCACGG	Mediterranean	↓ mRNA translation
	α1	Initiation codon	ACCATGG->ACCGTGG	Mediterranean	↓ mRNA translation
	–α	Initiation codon	ACCATGG->ACCGTGG	Black	↓ mRNA translation
	$-\alpha^{3.7II}$	Initiation codon	ACCATGG->AC-CATGG	Mediterranean, N.Africa	↓ mRNA translation
	α2	Exon III	α116GAC->UAG	Black	in phase termination
	α2	Termination codon	α142 TAA->CAA	Southeast Asean	Hb CS
	α2	Termination codon	α142 TAA->AAA	Mediterranean	Hb Icaria
	α2	Termination codon	α142 TAA->TCA	Indian	Hb Koya Dora
	α2	Termination codon	α142 TAA->GAA	Black	Hb Seal Rock
	–α	Exon I	α30/31GAGAGG->GAGG	Black	reading frameshift
Posttranslational	α2	Exon III	α125 Leu->Pro	Southeast Asean	Hb Quong Sze
instability	α2	Exon III	α109 Leu->Arg	Southeast Asean	Hb Suan Dok
	α	Exon III	α110 Ala->Asp	Middle Eastern	Hb Petah Tikvah
	–α	Exon I	α14 Trp->Arg	Black	Hb Evanston
Uncharacterized	α	Unknown	Not determined	Black	
	α	Unknown	Not determined	Greek	'Karditsa' mutation

Table 1. Nondeletional mutant that cause α-thalassemia (according to Higgs et al., 1989)

1.2.6 The clinical dyndromes of α - thalassemia

It is possible to subdivide thalassemic clinical syndromes into four categories : silent carrier, α - thalassemia trait, Hb H disease, and α - thalassemia hydrops fetalis. There can be significantly overlap among these groups based upon the spectrum of clinical severity which reflects the variety of genetic background. It may not always be possible to the basis of clinical data alone.

1. Silent Carrier (three functional α - genes): the hematological parameters such as hemoglobin concentration, red cell indices and number of red cells are within normal limit.
2. Alpha thalassemia trait (two functional α - genes): there are mild hematological changes but no major clinical abnormality.
3. Hb H disease (one functinal α - gene) : the clinical severity is considerable variability.
4. Hb Bart's hydrop fetalis(no functional α - gene) : the lost of all four α - globin genes is in compatible with life, resulting in mid- to late-gestational stillbirth of an hydropic fetus.

1.2.6.1 Silent carrier : the lost of a single α - globin gene

The $\alpha^{3.7}$ deletion is by far the most common single gene disorder worldwide (Hess, 1984). In some populations it is found in the majority of the population. The deletion a single α - globin gene is phenotypically silent . The hematologic parameters such as hemoglobin concentration, red cell indices, and number of red cells are within normal limits. Even by measuring the α / β globin chain synthetic ratio in reticulocytes, it is not possible to distinguish these individuals from normal. Individuals with the -α / αα genotype can be specifically diagnosed by directly detecting the deletion by DNA mapping or Southern blot hybridization (Leibhaber, 1989)

1.2.6.2 Alpha thalassemia trait : the lost of two α – globin genes

The most frequent cause of α – thalassemia trait is homozygosity for the $-\alpha^{3.7}$ deletion ($-\alpha^{3.7}$ / $-\alpha^{3.7}$). A second, less frequent -- / α α genotype can also underly this phenotype, most commonly in Southeast Asian populations. This genotype is usually associated with a significant microcytosis (MCV usually in the range of 70-80 fl), an elevated red cell count, and an appreciable imbalance in the α / β ratio . Elevated level of Hb Bart's (γ 4) are found in cord blood . Hemoglobin levels in the adult are usually normal or only slightly depressed and Hb H is not present . Homozygotes for the $-\alpha^{4.2}$ deletion ($-\alpha^{4.2}$ / $-\alpha^{4.2}$) appear to have amore marked phenotype than homozygotes for the $-\alpha^{3.7}$ deletion ($-\alpha^{3.7}$ / $-\alpha^{3.7}$) with compound heterozygotes ($-\alpha^{3.7}$ / $-\alpha^{4.2}$) intermediate in severity (Bowden, 1987). The milder phenotype of the $-\alpha^{3.7}$ deletion probably reflects the compensatory increase in the expression of the remaining α – globin gene which may not occur in the $-\alpha^{4.2}$ deletion (Liebhaber, 1989).

1.2.6.3 Hb H disease

Hb H disease is the most severe of the α-thalassemia phenotypes compatible with life. It most frequently results from the interaction of α - thalassemia 1 and α - thalassemia 2, and therefore it is predominantly found in Southeast Asia (commonly -- SEA / $-\alpha^{3.7}$) and the Mediterranean basin (commonly -- MED / $-\alpha^{3.7}$) where both α - thalassemia 1 and α - thalassemia 2 are common. Hb H disease may also result from the interaction of nondeletion mutations affecting the predominant α2 globin gene ($\alpha^{Nco}\alpha$ / $\alpha\ ^{Nco}\alpha$, $\alpha^{T\ SAYDI}$ / $\alpha\alpha^{TSAUDI}$, and $\alpha\alpha\alpha\ ^{TSAUDI}$ / $\alpha\alpha\ ^{TSAUDI}$). In Algeria, homozygotes for the $-\alpha^{3.7\ IIT}$ defect ($-\alpha^{3.7\ IIT}$ / $-\alpha^{3.7\ IIT}$) have typical Hb H disease (Whitelaw et al., 1986).

The clinical picture of Hb H disease is that of a chronic hemolytic anemia of variable severity. Hemoglobin levels are usually around 8-10 g / dl but the variations can be much wider. The anemia is accompanied by jaundice, hepatosplenomegaly and a

number of complications such as leg ulcers, gallstones and folic acid deficiency. Electrophoresis of the hemolysate shows the presence of HbH in the range 0.8-40% of the total hemoglobin and during growth is gradually but not totally replaced by HbH, so that it is occasionally found in the peripheral circulation of adult individuals. The synthesis of α chains in recticulocytes is reduced to about 25 % of the normal values. Examination of the blood smear reveals hypochromia, microcytosis, poikilocytosis, target cells and, after incubation of blood with brilliant cresyl blue, numerous red cells with HbH inclusion bodies. The number of these cells increases after splectomy and is associated with the presence of large HbH inclusion bodies. Chronic hemolysis is intensified during acute episodes owing to infections or administration of oxidizing drugs. In some cases a very mild, asymptomatic HbH disease has been revealed by the acute hemolysis following such events. In keeping with the lower α - globin chain production of the $\alpha^T\alpha$ haplotype, homozygotes for such a haplotype may present with a mild HbH disease (Fei , 1992) . On the other hand, the hemolytic anemia observed in patients with the genotype -- / $\alpha^T\alpha$ is usually more severe than that found in HbH disease caused by deletional α gene mutations (Galanello , 1992; Lui , 1994; Kanavakis , 1996). In rare cases the interaction of α^0 - thalassemia with a non- deletion α^+ haplotype has led to hydrops fetalis (Chan et al, 1997). The severity of this combination has motivated in a number of case the selective interruption of the pregnancy. It is important to notice, however, that in the case of interaction of the haplotype α^0 with Hb Icaria or Constant Spring the clinical course has been favorably modified by splenoctomy (Kanavakis et al, 1996).

In Thailand, where there is an abundance of well documented cases of Hb H disease, it is known that despite the relatively homogenous nature of the molecular basis (--SEA / $--\alpha^{3.7}$ in 80 %) (Winichagoon et al., 1984), the clinical course is quite variable. This suggests that other genetic and environmental factors play an important role in the clinical and hematologic variation seen in this syndrome.

1.2.6.4 Hb Bart's hydrops fetalis

Hb Bart's hydrops fetalis is the most severe form of α-thalassemia , resulting form a complex lack of α-chain production, which is almost always deletional in origin [--/--]. Lost of expression of all four α-globin genes, these infants are profoundly anemia and their red cells contain only Hb Bart's [γ_4] and Portland ($\zeta_2\gamma_2$). The physiological nonfunctional hemoglobin, Hb Bart's is produced by the association of chains to form homotetramers. Residual expression of the α- globin gene result in synthesis of sufficient functional hemoglobin tetramers [Hb portland] to carry the fetus through mid to late gestation. Affected infants either are still born between 34 and 40 weeks gestation or are still born alive but die within the first few hours. There is pallor, edema, and hepatosplenomegaly, and the clinical picture resembles hydrops fetalis due to Rh blood group incompatibility. At autopsy there is massive extramedullary hemopoeisis enlargement of the placenta. A variety of congenital anomalies have been observed. There have been a few reports of the rescue of infants with this syndrome by prenatal detection and exchange transfusion. These babies have grown and developed normally, although they are of course blood transfusion-dependent. This condition is associated with a high incidence of maternal toxemia of pregnancy and difficulties at the time of delivery because of the massive placenta. The reason for placental hypertrophy is unknown, although, because a similar phenomenon is observed in hydrops infants with Rh incompatibility, it may reflect severe intrauterine hypoxia (Weaterall,1995).

1.3 The molecular approaches in detection of α-thalassemia 2 ($-\alpha^{3.7}$)

1.3.1 PCR technology

The polymerase chain reaction (PCR) is a method that can amplify the selected fragment of DNA in vitro. Since it was introduced by Mullis at the Cetus Corporation in 1983, the technique has been modified for many uses and has essentially revolutionized molecular biology (Guyer and Koshland, 1990). The general principles and the optimization of the PCR below are summarized from "PCR Technology " : Principles and applications for" DNA Amplifications " by Henry A Erlich (Editor), 1992 and from "PCR Protocols : A Guide to Methods and Application ", Innis et al.,(1990).

1.3.2 General principles

The PCR technique was developed highly efficient amplification of DNA sequences of interest. In general, the procedures depends on the availability of sequences that flank region of interest. Two synthetic oligonucleotides are prepared using these flanking sequences, one complementary to each of the strand. The DNA is denatured at high temperature and then reannealled in the presence of large molar excess of the oligonucleotides. The oligonucleotides, oriented with their 3' ends pointing toward each other, hybridize to opposite strand of the target sequence and prime enzymatic extension along the nucleic template in the present of the deoxynucleotide triphosphates. The end product is then denatured again for another cycle. Since the product of one cycle can serve as templates for the next, the number of the product is increased exponentially as a function of cycle number. This leads to the selective enrichment of specific DNA sequences so that they can be readily manipulated or detected.

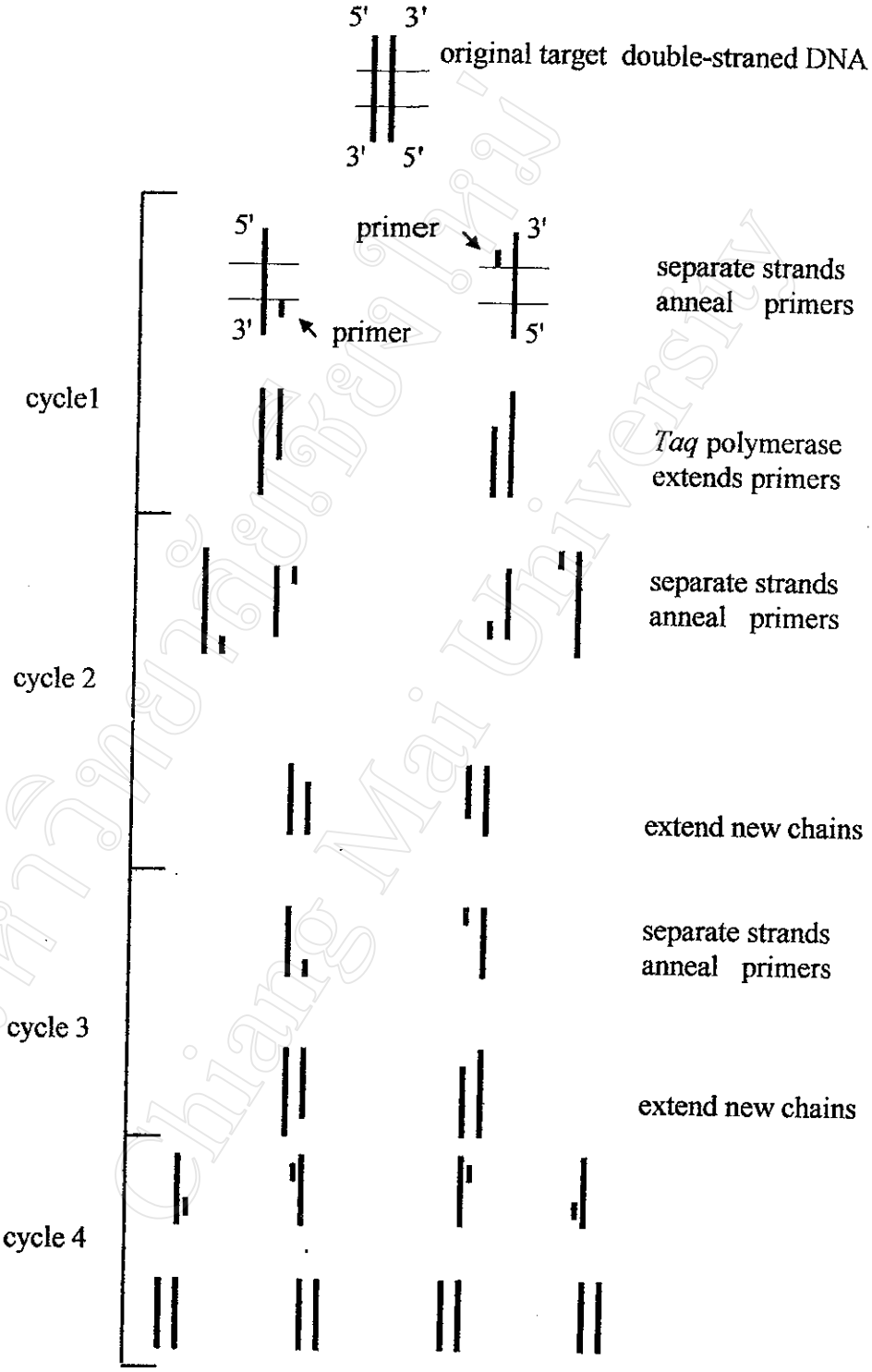

Figure 7. Schematic representation of amplification of target DNA sequence by Polymerse Chain Reaction technique .

1.3.3 The standard reaction

The standard PCR is typically done in a 50-100 µl volume and, in addition to the sample DNA, contains 50 mM KCL, 10 mM Tris-HCl (pH8.4), 1.5 mM $MgCl_2$, 100 µg/ml gelatin, 0.25 µm of each primer, 200 µM of each deoxynucleotide triphosphate, and 2.5 units fo *Taq* polymerase, The type of mineral oil are often added to seal the reaction and prevent condensation.

The amplification can be conveniently perform in a DNA Thermal Cycler (Perkin- elmer Cetus Instrument) using the " Step Cycle " program set to denature at 94° C for 1 minute, anneal at the proper temperature for a total of 30 cycles. These conditions can be use to amplify a wide range of target sequences with excellent specificity. However, some PCR application may need the optimizations, which are described below, for the best result.

Optimization strategies for the PCR

Individuals reaction components and time / temperature parameters must be adjusted for efficient amplification of specific targets.

Primer selection

There is no set of rules that will ensure the synthesis of an effective primer pair. However, the following guideline is useful for the designing of primers.

1. Where possible, select primers with a random base distribution and with a GC content similar to that of the fragment being amplified. Try to avoid primers with stretches of polypurines, polypyrimidines, or other unusual sequences.
2. Avoid sequences with significant secondary structure, particular at the 3' end of the primer.

3. Check the primers against each other for complementary. In particular, avoiding primers with 3' overlaps will reduce the incidence of "primer dimer". Primer dimer occurs when the 3' ends of the two primers hybridize, they form a "primer template" complex, and primer extension results in a short duplex product.

PCR primers are oligonucleotides, typically 15 to 30 bases long, and are complementary to sequences defining the 5' ends of the complementary template strands. Primer concentration between 0.1 and 1.0 μM are generally optimal. The use of higher concentrations of primer is not only expensive but also increases the chances of mispriming with the resultant formation of non-specific product. Using lower primer concentrations may also help reduce the primer dimer artifact.

Buffer and magnesium ion

Buffer usually provided along with the DNA polymerase enzyme in the form of 10x buffer. The company usually titrates for the optimal concentration of the components within the buffer; there for , it is no need to change the composition of the buffer unless the special effect such as reduction of the secondary structure is needed. The free magnesium ion concentration must be adjusted for specific PCR experiments. An optimal concentration usually between 1.5 and 4.5 mM. Too little free magnesium may produce a variety of unwanted products.

Deoxynucleotide triphosphates

In the standard protocol, each dNTP concentration is usually presented at 50-200 μM. Higher concentrations tend to promote misincorporations by the polymerase and should be avoided. Lower concentrations may give higher fidelity and specificity with no reduction in product yield.

DNA polymerase

Until now there are many thermostable DNA polymerases available, but *Taq* DNA polymerase is still be the enzyme of choice, due to the historical profile and widely distribution.

In the 100μl reaction volume, 2.0 units *Taq* DNA polymerase are recommended for most PCR applications. The enzyme can be added conveniently to a fresh master mix prepared for a number of reactions, therefore avoiding the tedious process and possible accuracy problems associated with adding individual 0.5 μl enzyme aliquots to each tube.

Target DNA

Target DNA should be in a low EDTA containing buffer since the magnesium ion is essential for the reaction . It is sometimes recommended to dissolve the target DNA in purified water. The target DNA concentration varies, depending on the type of subsequences to be amplified. For plasmid DNA , nanogram amounts are good staring point. For genomic DNA, amounts ranging from 0.05-1.0 μg are typically used for amplifications of single loci.

Cycling parameter

PCR is performed by incubating the samples at tree temperatures corresponding to the three step in a cycle of amplification: denaturation, annealing, and polymerization.

Typical denaturation conditions are 90-95 ^{0} C for 30-60 seconds. However, the higher temperature on the longer time may be appropriate, especially for GC-rich target. In contrast, too high or too long denaturation step lead to unnecessary loss of enzyme activity. The half- life of *Taq* DNA polymerase is > 2 hours at 92.5 ^{0} C, 40 minutes at 97.5 ^{0} C. Annealing is a very critical step for the specificity of the PCR products. The annealing temperature should be adjusted each reaction. Usually the temperature in the

range 55-72^0 C yield the best results. The estimate annealing temperature for each primer can be calculated by the following formula: 2 AT + 4 GC - 5 ^{0} C; or 5^0 C below the true Tm of the primers.

Polymerization usually set at 72^0 C, the time is usually depended on the length of the desired PCR products. At the optimum temperature of *Taq* DNA polymerase (75-80^0 C), the Kcat approaches 150 nucleotides/ second. At 72 ^{0} C, the rate of polymerization is around 2-4 kb/minute. Thus, 30 seconds usually enough for the extension of PCR products less than 1000 nucleotides.

มหาวิทยาลัยเชียงใหม่ Chiang Mai University

1.4 Objectives

1. To find the optimal conditions for the polymerase chain reaction technique in order to detect alpha thalassemia 2 carrier : rightward type (- $\alpha^{3.7}$)
2. To study the prevalence of alpha thalassemia 2 carrier : rightward type (- $\alpha^{3.7}$) in the Northern Thai population by polymerase chain reaction technique
3. To study the hematological characteristic of alpha thalassemia 2 carrier : rightward type (- $\alpha^{3.7}$) in the Northern Thai population.

มหาวิทยาลัยเชียงใหม่ Chiang Mai University