

MOLECULAR GENETICS OF β THALASSAEMIA

S.L. Thein

β Thalassaemia refers to a group of inherited haemoglobin disorders characterised by a reduced synthesis of β globin chains. It is found throughout the world, prevalent in the tropical and sub-tropical regions including the Mediterranean, Middle East, parts of Africa, Indian subcontinent, Southeast Asia and Southern China. It appears that heterozygotes for β thalassaemia are protected from the severe effects of falciparum malaria, and natural selection has increased and maintained their gene frequencies in these malarious regions.

The β thalassaemias are considered to be autosomal recessive disorders since individuals who have inherited one abnormal β gene (carrier) are clinically asymptomatic with minor haematological abnormalities, and the inheritance of two abnormal β globin genes (homozygotes or compound heterozygotes) is required to produce a clinically detectable phenotype.

Molecular Basis of β Thalassaemia

The β -like globin genes are arranged in a cluster on the short arm of chromosome 11 (11p15.5), in the order (5'- ϵ - γ^G - γ^A - ψ - β - δ - β -3') in which they are expressed during development. While the α -like genes undergo a single developmental "switch" (embryonic \rightarrow fetal/adult), the β -like genes undergo two "switches" (embryonic \rightarrow fetal \rightarrow adult). Mutations affecting the β globin gene only become apparent clinically on completion of the switch from fetal (Hb F, $\alpha_2\gamma_2$) to adult (Hb A₂; $\alpha_2\beta_2$) haemoglobin at two years of age.

Expression of the individual genes within the β cluster is controlled by complex interactions between the local regulatory sequences within each gene and the β locus control region (β LCR), the major regulatory elements upstream of the cluster. The importance of the various regulatory elements is illustrated by the striking heterogeneity of β thalassaemia genes caused by lesions in these regions. The latest repository of the mutations causing β thalassaemia in hemoglobin shows that > 150 β thalassaemia alleles have now been characterised.

In contrast to α thalassaemia, the vast majority of mutations causing β thalassaemia are non-deletional due to single base substitutions, small insertions or deletions (see Table 1).

Due to the vast number of different β thalassaemia mutations many patients with thalassaemia major are compound heterozygotes for two different molecular lesions.

Non-Deletion Forms of β Thalassaemia

These point mutations involve the critical sequences that control the various stages of gene expression and provide a good example of the spectrum of naturally occurring lesions that can inactivate a mammalian gene.

Approximately half of these β thalassaemia alleles completely inactivate the β gene and cause β^0 thalassaemia. Most are caused by the introduction of a premature termination codon due to frameshift or nonsense mutations. Low levels of nuclear and cytoplasmic mutant β globin mRNA are found in red cell precursors in these mutations. It is not clear how a premature termination codon could lead to a reduction of mRNA.

Table 1. Mutations Causing β Thalassaemia

Deletional	No.	Type
	17	
Upstream deletions	3	β^0
b gene deletions	4	β^0
	1	
Nondeletional	No.	Type
	130	
Transcriptional Mutants	18	Mild to very mild β^+
5' untranslated region (5'UTR)	5	very mild β^+
RNA Processing Mutants	38	
Splice Junction	15	β^0
Consensus site	13	Severe to mild β^+
Aberrant splicing	10	β^+
Introns	6	β^+, β^0
Coding regions	4	β^+
RNA Translation mutants	56	
Nonsense	12	β^0
Frameshift	44	β^0
RNA Cleavage and polyadenylation	6	mild β^+
CAP (+1) site	1	very mild β^+
3' untranslated region (3'UTR)	1	?
Initiation codon	5	β^0
Unlinked to b cluster	No.	Type
	3	β^+
Total	150	

Studies show that the different in-phase termination mutants exhibit a 'positional' effect. Frameshifts and nonsense mutations that result in premature termination early in the sequence (in exon 1 and 2) are associated with minimal amounts of mutant β mRNA.

In heterozygotes for such cases, no β chain is produced from the mutant allele and only half the normal β globin is present, resulting in a typical asymptomatic phenotype. In contrast, mutations that produce in-phase terminations later in the β sequence, in exon 3, are associated with comparable amounts of mutant β -mRNA. Such mutations even when present in a single copy, results in a moderately severe anaemia and are said to be 'dominantly inherited'. Small amounts of truncated β variant chains have been isolated in one case (heterozygous β codon 121). These truncated β chains, however, are non-functional and are not able to form viable tetramers, thus resulting in ineffective erythropoiesis and clinical disease even in the heterozygous state.

Other mutations which result in β^0 thalassaemia include those which affect the initiation codon (ATG) and mutations at the splice junctions (5' and 3') which completely abolish normal splicing.

The β^+ thalassaemia mutations allow the production of some β globin but the output is reduced. The reduction in β globin output ranges from minimal to almost complete absence. The severity of these β^+ thalassaemia alleles can be correlated to the degree of reduction in MCV in heterozygotes. A large number of β^+ thalassaemia alleles are caused by mutations that affect RNA processing and are located within the consensus sequence flanking the splice junctions or within introns or exons. The latter lead to the creation of new splice sites, which can partially or completely eliminate normal splicing. An example of a cryptic splice site involved in alternative splicing can be found in exon 1 of the β gene, at codons 24-27. A mutation within this region can make the site resemble more closely a true splice site. An important example is the G \rightarrow A (Glu \rightarrow Lys) mutation in codon 26 which leads to the production of Hb E. There is a low level of normally spliced mRNA which contains the exon 1 mutation leading to the production of Hb E as well as abnormal splicing into the codons 24-27 region which does not produce any recognisable β globin. The overall reduction in splicing is the molecular basis for the mild β^+ thalassaemia phenotype of Hb E. The β^E gene is prevalent in Southeast Asia, reaching a frequency of 75% in north east Thailand. Its interaction with β thalassaemia accounts for a large proportion of the thalassaemia major in Southeast Asia. Mutations affecting the conserved sequences in the 5' promoter, i.e., TATA box, proximal CACCC and distal CACCC box, typically cause a 70-80% reduction in promoter activity and are often very mild. Mutations affecting the polyadenylation signal (AATAAA) at the 3' end also generally result in a mild β^+ thalassaemia phenotype.

Among the β^+ thalassaemia alleles is a sub-group which cause a minimal deficit in β chain production. Heterozygotes for such mutations have normal Hb A2 levels and normal red cell indices and are often referred to as 'silent' carriers. This group includes the C-T mutation at position -101 upstream of the β gene, the A \rightarrow C mutation at position CAP (+1), and the mutations in the 5' untranslated region. The β thalassaemia is 'silent' when present in a single copy but becomes evident in homozygotes or compound heterozygotes for one of these mutations.

Deletions Causing β Thalassaemia

β thalassaemia is rarely caused by deletions; 17 deletions affecting the β cluster and causing β thalassaemia have been observed. These deletions fall into two classes: upstream deletions and deletions involving the β gene.

Upstream deletions

Three deletions, described in families of Dutch, English and Hispanic origin, are of particular interest because they remove substantial regions of the 5' end of the β gene clusters but leave the β gene itself intact and yet result in a β thalassaemia phenotype. These deletions silence the β globin gene because they remove all or a substantial proportion of the regulatory sequences in the β LCR. While deletion of the three 5' LCR elements (hypersensitive sites 2-4) inactivates the β gene, a family study showed that deletion of the 3' most LCR element (HS1) does not affect the activity of the β gene.

Deletions Involving the β Globin Gene

Fourteen different deletions affecting only the β globin gene have been described. These deletions range from 290 bp to > 45 kb. Of these, only the 619 bp deletion at the 3' end of the β gene is common, but even that is restricted to the Sind populations of India and Pakistan where it constitutes ~20% of the β thalassaemia alleles. The other deletions, although extremely rare, are of particular phenotypic interest because they are associated with an unusually high level of Hb A₂ in heterozygotes. The mechanism underlying the markedly elevated levels of Hb A₂ and the variable increases in Hb F in heterozygotes for these deletions is related to the removal of the 5' promoter region of the β globin gene, which removes competition for the upstream β -LCR leading to an increased interaction of the LCR with the γ - and δ -genes in cis, thus enhancing their expression.

Other deletions which lead to β thalassaemia remove extensive regions of the cluster including the ϵ , γ^G , γ^A , $\psi\beta$ and δ genes and part or all of the β gene.

Phenotypic Variants of β Thalassaemia

Dominantly Inherited β Thalassaemias

The dominantly inherited β thalassaemias produce a clinically detectable phenotype when present in a single copy, whereas individuals heterozygous for β thalassaemia are typically clinically asymptomatic. This group of variant β thalassaemias includes the β thalassaemic haemoglobinopathies, which can be defined as structural haemoglobin mutants presenting with a phenotype of thalassaemia. The underlying molecular abnormalities are remarkably heterogeneous and include amino acid substitutions, amino acid deletions, and elongated or truncated β globin chains (Table 2). Although nucleotide sequence analysis predicts the synthesis of a β variant, it is unusual to demonstrate an abnormal β globin from these mutations. Globin biosynthesis studies including short-term incubations and pulse chase experiments showed a rapid decline of some of the variants, suggesting that these variants are highly unstable and rapidly catabolised.

Thus, unlike the recessive forms of β thalassaemia, the dominantly inherited β thalassaemias are characterised by the synthesis of a highly unstable β variant. The phenotype resembles the intermediate forms of β thalassaemia due to ineffective

erythropoiesis, while it also bears resemblance to the congenital haemolytic anaemias due to the variable amount of peripheral haemolysis.

The phenotype of this class of β thalassaemia appears to be dependent on the extent of ineffective erythropoiesis and intravascular haemolysis, which in turn depends on the stability of the β chain variant, its ability to form α/β dimers and tetramers, and the stability of the tetramers.

Unlike typical β thalassaemia which is prevalent in malaria-endemic regions, dominantly inherited β thalassaemias are rare, occurring in dispersed geographical regions, including Northern and Eastern Europe, Japan and Korea, where the gene frequency for β thalassaemia is very low. Except for the codon 121 GAA \rightarrow TAA mutation, all the dominant β thalassaemia alleles have been described in single families, many as de novo events (Table 2). It is postulated that the low frequency of the dominant β thalassaemia mutations is due to the lack of positive selection as in the case for the recessive forms.

Clinically, since spontaneous mutations are common in dominant β thalassaemia, it is important that the disorder should be suspected in any patient with a thalassaemia intermedia phenotype even if both parents are haematologically normal and the patient is from an ethnic background where β thalassaemia is rare.

b Thalassaemia Due to a trans-Acting Determinant

It has been estimated that ~1% of the β thalassaemia genes in the world remain uncharacterised. In such cases, it has been postulated that mutations may be found in the upstream β -LCR or in the 3 enhancer region from 700-1100 nucleotides 3' of the β gene. Three families have been described in which the genetic determinant responsible for the β -thalassaemia phenotype segregates independently of the β globin gene cluster implicating the presence of a trans-acting determinant.

Intermediate Forms of β Thalassaemia

Thalassaemia intermedia is an ill-defined clinical term used to describe patients with phenotypes that are more severe than the asymptomatic thalassaemia trait but milder than the transfusion-dependent thalassaemia major. The criteria on which the diagnosis is based are that patients present later in life relative to thalassaemia major and that they are capable of maintaining a reasonable level of haemoglobin (3-6 gm/dl) without transfusion. At the severe end of the spectrum, patients present between the ages of 2 and 6 years, and although they are just capable of surviving without blood transfusion, it is clear that growth and development are retarded. Many will show the skeletal and facial changes and progressive splenomegaly as seen in untreated thalassaemia major. As they become older they develop iron overload because of increased gastrointestinal absorption of iron. At the other end of the spectrum, patients are completely asymptomatic until adult life and are transfusion independent, with haemoglobin levels of 10-12 gm/dl. Such patients are diagnosed either during episodes of infection when they become anaemic or by a chance haematological examination. There is usually some degree of splenomegaly.

Table 3 lists some of the molecular interactions associated with the phenotype of thalassaemia intermedia.

Because of the extreme variability of these disorders, these patients should be regularly followed from early childhood and the disease carefully monitored in terms of the growth charts and iron accumulation.

Table 2. Dominantly Inherited β Thalassaemias (including hyperunstable Hbs).

	β variant	Ethnic Group
I Single base substitutions		
i)	Cod 28 (CTG→CGG)Hb Chesterfield* Leu→Arg	English
ii)	Cod 30 (AGG→ACG) Hb Monroe Arg→Thr	Black American
iii)	Cod 32 (CTG→CAG)Hb Medicine Lake* Leu→Glu in <i>cis</i> with Cod 98 (GTG→ATG) Val to Met	Caucasian American
iv)	Cod 60 (GTG→GAG) Hb Cagliari* Val to Glu	Italian
v)	Cod 110 (CTG→CCG) Leu to Pro	Hb Showa-Yakushiji Japanese
vi)	Cod 114 (CTG→CCG) Leu to Pro	Hb Durham Irish American
vii)	Cod 115 (GCC→GAC) Ala to Asp(Hb HK)	Hb Hradec Kralove Czech
viii)	Cod 127 (CAG→CCG) Gln to Pro	Hb Houston N. Italian / French
II Deletion of Intact Codons → Destabilisation		
I)	Cod 32/34 (-GGT) Val-Val to Val	Hb Korea* Korean
ii)	Cod 127/128 (-AAG) Hb Gunma Glu-Ala to Pro	Japanese
iii)	Cod 134-137 (-12, +6) Val-Ala-Gly-Val to Gly-Arg	Portuguese
III Premature Termination → Truncated β Variant		
i)	Cod 121 (GAA→TAA) Glu to Term	Several families ** Caucasian, N. Europeans
ii)	Cod 127 (CAG→TAG) Gln to Term	English

IV Frameshifts → Elongated β Variants

I)	Cod 94 (+TG) → 156aa	Hb Agnana*	S. Italian
ii)	Cod 109 (-G) → 156aa	Hb Manhattan	Askenazi Jew
iii)	Cod 114 (-CT, +G) → 156aa	Hb Geneva	Swiss-French
iv)	Cod 123 (-A) → 156aa	Hb Makabe	Japanese
v)	Cod 123-125 (-ACCCACC)		Thai
vi)	Cod 124 (-A)		Russian
vii)	Cod 125 (-A)		Japanese
viii)	Cod 126 (-T) - 156aa	Hb Vercelli*	N. Italian
ix)	Cod 128/129 (-4, +5, -11) → 153aa		Irish

* Spontaneous mutations

** Several families reported including one spontaneous mutation

Distribution of β Thalassaemia in Different Populations

Molecular analysis of the β thalassaemia genes has demonstrated a striking heterogeneity, yet population studies indicate that probably only 20 β thalassaemia alleles account for > 80% of the β thalassaemia mutations in the whole world. This is due to the phenomenon of geographical clustering where each population has a few (4-6) common mutations together with a varying number of rare ones. This is particularly relevant to prenatal diagnosis because direct detection of these mutations by DNA analysis becomes feasible. Although the same mutation can occur in different populations, it is normally associated with a different β haplotype, implying that the mutation has arisen independently and of recent origin. Presumably these genes then achieve their high frequency through positive selection from malaria. In this respect it is interesting to compare the spectrum of β thalassaemia alleles in the high frequency areas with that in regions where β thalassaemia is uncommon, such as Northern Europe and Japan. In regions where β thalassaemia is prevalent, a limited number of alleles (4-5) account for 90% or more of the β thalassaemia, while in countries relatively free from malaria and where β thalassaemia is uncommon, a diversity of β thalassaemia mutations has been found. A study of β thalassaemia in indigenous Britons identified nine alleles, of which six were of the typical recessive type which accounted for 15 of the 20 β thalassaemia genes. The other three alleles, of which two were novel, were associated with a

Table 3. Molecular Basis of β Thalassaemia Intermedia

- I Homozygous or compound heterozygous state for β thalassaemia
 1. inheritance of mild β^+ thalassaemia alleles e.g. β IVS1-6 T-C, β promoter mutations

2. co-inheritance of α thalassaemia
 - effect more evident in β^+ thalassaemia
 3. β thalassaemia with elevated γ chain production
 - polymorphism at position -158 $^G\gamma$ gene (*Xmn* I- $^G\gamma$ site)
 - γ promoter mutations
 - heterocellular HPFH
X-linked, 6q-linked
- II Compound heterozygotes for β thalassaemia and deletion forms of HPFH or $\delta\beta$ thalassaemia
- III Compound heterozygotes for β thalassaemia and β chain variants, e.g. Hb E/ β thalassaemia
- IV Inheritance of deletion forms of β thalassaemia which remove the 5' β promoter region
- V Heterozygotes for β thalassaemia
1. co-inheritance of extra α globin genes ($\alpha\alpha\alpha/\alpha\alpha$ or $\alpha\alpha\alpha/\alpha\alpha\alpha$)
 2. dominantly inherited forms of β thalassaemia (including some thalassaemic haemoglobinopathies)

dominant phenotype. A similar diversity of β thalassaemia has also been found in Japan where there is a large proportion of isolated mutations that were dominantly inherited as well as recessive β thalassaemia mutations, many of which are unique to Japan. This diversity of β thalassaemia mutations is probably typical of populations from non-malarial, temperate regions where there is no selective pressure from malaria.

References

1. Baysal E, Carver MFH: The β - and δ -thalassemia repository. *Hemoglobin* 19:213, 1995
2. Bunn HF, Forget BG: *Hemoglobin: Molecular, Genetic and Clinical Aspects*. Philadelphia, PA, W.B. Saunders Company, 1986
3. Crossley M, Orkin SH: Regulation of the β -globin locus. *Curr Opin Genet Dev* 3:232, 1993
4. Grosveld F, Dillon N, Higgs D: The regulation of human globin gene expression, in Higgs DR, Weatherall DJ (eds): *Baillière's Clinical Haematology*, vol 6:1. London, Baillière Tindall, 1993, p 31
5. Olivieri NF, Brittenham GM, Matsui D, Berkovitch M, Blendis LM, Cameron RG, McClelland RA, Liu PP, Templeton DM, Koren G: Iron-chelation therapy with oral deferiprone in patients with thalassaemia major. *N Eng J Med* 332:918, 1995
6. Rodgers GP, Rachmilewitz EA: Novel treatment options in the severe β -globin disorders. *Brit J Haemat* 91:263, 1995

7. Stamatoyannopoulos G, Nienhuis AW, Majerus PW, Varmus E: The Molecular Basis of Blood Diseases. Philadelphia, Saunders, 1994
8. Thein SL: Dominant β thalassaemia: molecular basis and pathophysiology. *Brit J Haemat* 80:273, 1992
9. Thein SL: β -thalassaemia, in Higgs DR, Weatherall DJ (eds): *Baillière's Clinical Haematology. International Practice and Research: The Haemoglobinopathies*, vol 6:1. London, Baillière Tindall, 1993, p 151
10. Weatherall DJ, Clegg JB: *The Thalassaemia Syndromes*. Oxford, Blackwell Scientific, 1981
11. Hall GW, Barnetson RA, Thein SL: Beta thalassaemia in the indigenous British population. *Brit J Haemat* 82:584, 1992
12. Hall GW, Thein SL: Nonsense codon mutations in the terminal exon of the β -globin gene are not associated with a reduction in b-mRNA accumulation: a mechanism for the phenotype of dominant β -thalassemia. *Blood* 83:2031, 1994
13. Thein SL, Wood WG, Wickramasinghe SN, Galvin MC: β -thalassemia unlinked to the β -globin gene in an English family. *Blood* 82:961, 1993
14. Murru S, Loudianos G, Porcu S, Sciarratta GV, Agosti S, Parodi MI, Cao A, Pirastu M: A β -thalassaemia phenotype not linked to the β -globin cluster in an Italian family. *Brit J Haemat* 81:283, 1992
15. Pacheco P, Peres MJ, Faustino P, Pischedda C, Gonçalves J, Carvajales-Ramos M, Seixas T, Martins MC, Moi P, Lavinha J: β -thalassaemia unlinked to the β -globin gene interacts with sickle-cell trait in a Portuguese family. *Brit J Haemat* 91:85, 1995
16. Craig JE, Barnetson RA, Prior J, Raven JL, Thein SL: Rapid detection of deletions causing $\delta\beta$ thalassaemia and hereditary persistence of fetal hemoglobin by enzymatic amplification. *Blood* 83:1673, 1994