

# \* CHAPTER 10

## Molecular basis of thalassaemia syndromes

Bill Wood, Doug Higgs

## 1. Introduction

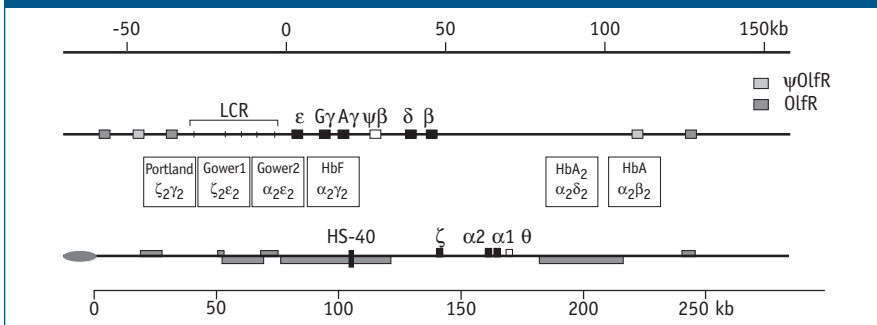
Taken together, the thalassaemias are the commonest human genetic disorders, with estimated carrier numbers of > 270 million and over 350,000 people affected with severe forms of the disease worldwide (1). It is most frequent in a band across the Old World that includes south-eastern and south Asia, the Middle East, the Mediterranean and north and central Africa. The high frequencies in these areas are believed to result from the relative protection of the heterozygotes against *Plasmodium falciparum malaria*. Increased emigration has meant that the diseases are now encountered worldwide.

The thalassaemias are characterised by deficient or absent production of one or other of the globin chains of haemoglobin. There are several  $\alpha$ -like and  $\beta$ -like globin genes that are developmentally regulated to produce embryonic ( $\zeta_2\varepsilon_2$ , HbGower1,  $\zeta_2\gamma_2$ , HbPortland,  $\alpha_2\varepsilon_2$  HbGower2), foetal ( $\alpha_2\gamma_2$ ,  $\alpha_2A\gamma_2$ ) and adult ( $\alpha_2\beta_2$  HbA,  $\alpha_2\delta_2$  HbA<sub>2</sub>) haemoglobins. Normal adult red cells contain 97% HbA with ~2.5% of the minor component HbA<sub>2</sub> and a small amount (< 1%) of HbF. A deficiency of  $\alpha$  chains causes  $\alpha$  thalassaemia while  $\beta$  thalassaemia has reduced  $\beta$  chain production. The molecular bases of these conditions has been well characterised and they have provided the paradigmatic examples for human genetics.

## 2. Structure of the globin gene clusters

The globin genes are encoded in separate gene clusters; the  $\alpha$  cluster lies at the telomere of chromosome 16 while the  $\beta$  cluster lies at chromosome 11p15.5. In each case, the genes are aligned 5' to 3' in the order in which they are expressed developmentally and both sets of genes are under the regulation of enhancer-like elements that lie some distance away at the 5' end of the cluster (Figure 1). Deletion

**Figure 1: The organisation of the  $\alpha$  and  $\beta$  globin gene clusters together with the composition of the haemoglobins produced at different developmental stages**



of these enhancer elements results in inactivation of any globin gene that remains structurally intact.

The  $\alpha$  cluster is surrounded by widely expressed “housekeeping” genes that lack tissue specificity and maintain an early replicating, open chromatin structure in both erythroid and non-erythroid cells. There are several upstream DNase1 hypersensitive sites (HSs), four of which are erythroid specific. Only one of these sites, 40kb upstream from the  $\zeta$  globin gene, has strong enhancer activity on its own. Known as HS-40, this site is believed to be the major  $\alpha$  globin gene regulatory element. It lies in an intron of a neighbouring widely expressed gene (C16orf35) of unknown function. The embryonic  $\zeta$  gene lies 5' of the duplicated  $\alpha$  genes. The  $\alpha$  genes differ by a few nucleotides in intron 2 and the 3'UTR but produce identical protein products. The output of the 5'  $\alpha 2$  gene exceeds that of the  $\alpha 1$  gene by 2-3 times. The  $\alpha$  cluster contains a pseudozeta and a pseudoalpha gene as well as two other transcribed genes,  $\alpha^D$  and  $\theta$ ; it is not known whether these are translated into protein products.

The  $\beta$  cluster is surrounded by numerous olfactory receptor genes. In non-erythroid cells the locus is in a closed chromatin conformation and is late replicating. Its ‘enhancer’ consists of five elements marked by erythroid-specific HSs lying 6-20 kb upstream of the  $\epsilon$  globin gene, each of which contains several binding sites for erythroid-specific and other transcription factors. Collectively these elements are known as the locus control region (LCR) and each contributes part of the overall LCR activity. In addition there is an erythroid specific hypersensitive site ~20kb downstream of the  $\beta$  globin gene. When the cluster is activated the upstream and downstream HSs are brought into close proximity with the gene promoters in a structure that has been dubbed an active chromatin hub (2). The mechanism by which the hub forms remains to be determined, although looping and tracking models have been suggested. The  $\beta$ -like globin genes are lined up in the order 5'- $\epsilon$ - $\gamma$ - $\delta$ - $\beta$ -3', with a pseudogene lying between the  $\gamma$  and  $\delta$  genes.

## 2.1 Structure of the globin genes

Globin genes consist of three coding sequence exons separated by two introns to give a total length of ~1500 nucleotides, a structure that has been highly conserved through evolution. Upstream of the first exon lies the promoter containing sequence motifs essential for specifying correct transcriptional initiation. A TATA box is found ~30bp upstream of the initiation site together with one or more CCAAT sites at ~70bp upstream. The promoters also contain a CACCC or CCGCCC box that binds EKLf1 and some have binding sites for the erythroid transcription factor GATA-1. The primary transcript contains a modified m<sup>7</sup>G CAP site, a 5' untranslated region (5'UTR), coding sequences separated by intron-exon junctions and a 3'UTR that

contains the poly A signal sequence, AAUAAA. Cleavage of the transcript occurs 20bp downstream of this signal and a 50-75 nucleotide polyA tail is added. Splicing out of the introns follows the canonical 5'GT-AG3' rule with the breakpoints surrounded by consensus sequences essential for proper splicing.

### 3. The $\alpha$ thalassaemias

#### 3.1 Phenotypes

Most of the mild forms of  $\alpha$  thalassaemia, caused by the loss of output of one or two of the four  $\alpha$  globin genes, are asymptomatic. The "silent" carrier state (loss of one gene) has minimal reductions in red cell size and haemoglobin content but may be detected by increased amounts of Hb Bart's ( $\gamma_4$ ) in some of the cases at birth. Loss of two genes results in mild anaemia with reduced red cell indices (MCV  $\sim$ 70fl, MCH  $\sim$ 22pg) and a normal pattern of haemoglobin analysis in adults. A similar phenotype is seen whether there is loss of one gene from each chromosome or loss of two genes from the same chromosome.

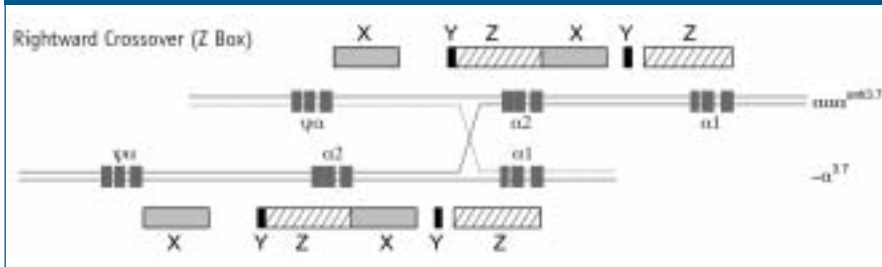
A loss of output from the equivalent of three genes results in a condition known as HbH ( $\beta_4$ ) disease, because the excess  $\beta$  chains tetramerise and can be detected in the red cells as 'golf ball' inclusions after incubation with supravital dyes. HbH disease results in a moderate to severe anaemia that may occasionally become transfusion dependent. A complete or almost complete loss of  $\alpha$  chain production results in Hb Bart's hydrops foetalis. This condition usually results in foetal death in the third trimester though early diagnosis and intrauterine transfusions have resulted in live births (3-6).

#### 3.2 The molecular basis of $\alpha$ thalassaemia

The  $\alpha$  thalassaemias are divided into those in which output from the affected chromosome is reduced ( $\alpha^+$ ) and those in which it is absent ( $\alpha^0$ ). The most common causes are deletions of one or both  $\alpha$  genes which are designated  $-\alpha$  and  $--$  respectively. The single  $\alpha$  gene state is believed to have arisen several times independently by crossover between two misaligned  $\alpha$  genes on the homologous chromosomes (Figure 2). Depending on the point of crossover, deletions may remove between 2.5 and 5.3 kb of sequence with the loss of 3.7 ( $-\alpha^{3.7}$ ) or 4.2 ( $-\alpha^{4.2}$ ) kb being the most prevalent. The remaining  $\alpha$  globin gene is expressed at a level between that of a normal  $\alpha 1$  and  $\alpha 2$  gene irrespective of which gene has been deleted.

The loss of both  $\alpha$  genes occurs as a result of illegitimate or non-homologous recombination, often involving Alu repeats, with lengths between 5.2 and  $>$  40kb. More than 20 different deletions have been described with the commonest being

**Figure 2: Misaligned crossing over in the  $\alpha$  cluster**



Misaligned crossover within the Z box (striped block) of the  $\alpha$  cluster to generate the common deletion form of  $\alpha$  thalassaemia  $\alpha^{3.7}$  and a  $\alpha\alpha\alpha$  chromosome. Crossovers within the homologous X boxes (shaded block) give rise to the  $-\alpha^{4.2}$  form of thalassaemia.

those from south east Asia, the Mediterranean and the Philippines, designated --SEA, --MED and --FIL respectively. The remainder either have a localised geographic distribution or are restricted to individual families. The longer deletions also remove the  $\zeta$  globin gene but this does not appear to compromise the survival of heterozygotes but is likely to be embryonic lethal in the homozygous state.

Rare deletions remove the upstream regulatory elements, leaving the structural globin genes intact but unexpressed. These are restricted to isolated families and are of variable length with several extending to the subtelomeric region. They all include HS-40 as well as one or more of the other erythroid hypersensitive sites in the upstream regulatory region and the minimal region of overlap extends ~20kb around this element. They are important in demonstrating the absolute requirement for this region on  $\alpha$  globin gene expression but, as yet, no natural deletion of the HS-40 element alone has been reported (7).

Several non-deletion types of  $\alpha$  thalassaemia (designated  $\alpha^T$ ) are known but they are much less common than the deletion forms and have a more restricted distribution. This is in contrast to the  $\beta$  thalassaemias where point mutations are the commonest types and gene deletions are rare. Mutations may occur in either the  $\alpha 1$  or  $\alpha 2$  genes with the latter appearing to be more common. This is probably an ascertainment bias as  $\alpha 2$  mutants are more likely to result in a detectable phenotype. The mutations may affect the initiation codon, splicing signals, cause frame shifts or introduce premature stop codons, a similar array to those causing  $\beta$  thalassaemia. Among the non-deletion  $\alpha$  thalassaemias are several that are due to missense mutations in which the amino acid substitution results in a highly unstable  $\alpha$  globin chain that is present in low amounts or is even undetectable by conventional protein analysis. Four mutations affect the  $\alpha 2$  polyA addition signal,

probably causing extended transcription, reduced amounts of normal message and possibly reading through and interfering with transcription of the downstream  $\alpha 1$  gene. One of these mutations ( $\alpha^{\text{TSaudi}}$ ) $\alpha$  is not an uncommon type in Saudi Arabia. Finally, at least five different nucleotide substitutions are known that convert the  $\alpha 2$  termination codon (TAA) into an amino acid coding triplet allowing translation to continue into the 3' UTR and producing an elongated  $\alpha$  chain containing an additional 31 amino acids. One of these mutants, Hb Constant Spring ( $\alpha^{\text{CS}}$ ) $\alpha$ , is not infrequent in south east Asia. It is believed that not only is the elongated  $\alpha$  chain unstable but that translation of the UTR affects the mRNA stability so that there is only ~1% Hb Constant Spring in the blood.

## 4. The $\beta$ thalassaemias

### 4.1 Phenotypes

The  $\beta$  thalassaemias are characterised by a reduction in output of  $\beta$  globin chains of variable degree. A complete absence of  $\beta$  production is known as  $\beta^0$  thalassaemia while a reduced amount of normal  $\beta$  chain production is found in the  $\beta^+$  thalassaemias. Very mild reductions are classed as  $\beta^{++}$  thalassaemia.

Heterozygotes for  $\beta^0$  or  $\beta^+$  have a classical thalassaemic haematology picture, with mild anaemia, microcytosis (MCV 70-80fl), hypochromia (MCH 18-22pg) and a blood picture of anisopoikilocytosis with target cells and basophilic stippling. Haemoglobin analysis shows an increased level of HbA<sub>2</sub> (3.5-6%) and normal or slightly raised (0.5-3%) levels of HbF.

Homozygotes for the more severe forms of  $\beta$  thalassaemia result in thalassaemia major. They present in early childhood with severe anaemia, hepatosplenomegaly and become transfusion dependent. Repeated transfusions and increased iron absorption lead to iron deposition in many organs (liver, endocrine tissues, heart) and ultimately death unless chelation therapy is used. Haemoglobin analysis shows mainly HbF plus a small amount of HbA<sub>2</sub> in  $\beta^0$  homozygotes, but insufficient to compensate for the loss of HbA, and the Hb level remains < 5 g/dL. Most  $\beta^+$  homozygotes and  $\beta^+\beta^0$  compound heterozygotes also have thalassaemia major. HbF remains the predominant haemoglobin but with variable amounts of HbA (5-30%), depending on the specific  $\beta^+$  thalassaemia allele.

Thalassaemia intermedia is used to describe cases with a more moderate clinical state in which transfusions begin later and may be required at less frequent intervals. Hb levels may stabilise at 6-9 g/dL without transfusions but will be exacerbated by hypersplenism or intercurrent infections. As the red cell pathophysiology in  $\beta$  thalassaemia is largely due to the effects of the excess  $\alpha$

chains, any genetic modifier that reduces this excess is likely to ameliorate the disease and be more likely to result in thalassaemia intermedia. Some of the commonest causes are given in Table 1. They include the presence of one or more mild  $\beta$  thalassaemia alleles, double heterozygotes for  $\alpha$  and  $\beta$  thalassaemia and conditions that lead to increased  $\gamma$  chain production (e.g. HPFH) which mops up some of the  $\alpha$  chain excess (8).

#### 4.2 Molecular basis of the $\beta$ thalassaemias

Over 200 different thalassaemic alleles of the  $\beta$  globin gene have been reported; the vast majority are caused by point mutations within the gene or its immediate flanking sequences. The distribution of alleles is highly variable from one population to another but within each population there are only a few alleles that are common (Figure 3). While most alleles behave as Mendelian recessives, there are variants that cause a disease phenotype even when present in a single copy. These act in a dominant negative fashion and are referred to as “dominantly inherited  $\beta$  thalassaemias”.

While simple deletions of the  $\beta$  globin gene are rare, there are more extensive deletions that remove both  $\delta$  and  $\beta$  genes and that may also extend to include the

**Table 1: Causes of thalassaemia intermedia**

##### 1. Mild deficit in $\beta$ chain production

Homozygosity for silent  $\beta$  thal  
 Homozygosity for mild  $\beta^+$  thal  
 Compound heterozygosity for severe  $\beta^0$  or  $\beta^+$  and silent or mild  $\beta^+$  thal  
 HbE/ $\beta$  thal  
 Heterozygous  $\beta$  thal with  $\alpha\alpha\alpha$  or  $\alpha\alpha\alpha\alpha$

##### 2. Reduced $\alpha$ chain excess

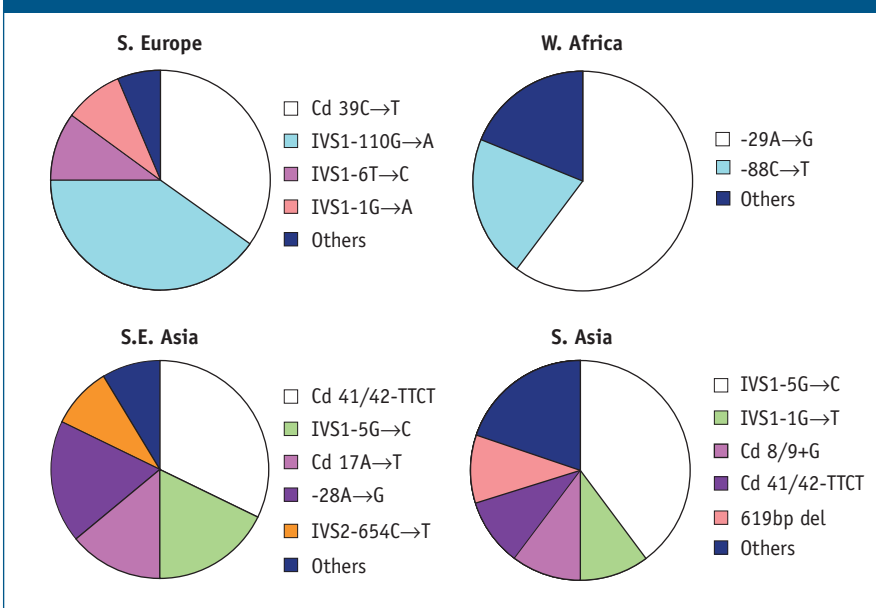
Homozygous  $\beta^0$  or  $\beta^+$  thal with 2 or 3  $\alpha$  gene deletions  
 Homozygous  $\beta^0$  or  $\beta^+$  thal with non-deletion  $\alpha$  thal  
 Homozygous  $\beta^+$  thal with 1 or 2  $\alpha$  gene deletions

##### 3. Increased $\gamma$ chain synthesis

Homozygous  $(\delta\beta)^0$  thal or  $(\gamma\delta\beta)^0$  thal  
 Compound heterozygosity for severe  $\beta^0$  or  $\beta^+$  and  $(\delta\beta)^0$  thal or HPFH  
 Compound heterozygosity for severe  $\beta^0$  or  $\beta^+$  and heterocellular HPFH  
 Homozygosity for Hb Lepore or Hb Lepore/ $\beta$  thal compound heterozygosity

##### 4. Dominant $\beta$ thalassaemia

Highly unstable  $\beta$  chain variants  
 $\beta$  thal heterozygotes with third exon mutations

Figure 3: The distribution of  $\beta$  thalassaemia alleles in various populations

The distribution of  $\beta$  thalassaemia alleles in various populations, illustrating the predominance of just a few alleles in each but which differ from one population to the other.

$\Delta\gamma$  gene. The lack of  $\delta$  and  $\beta$  chain production in these conditions is partially compensated, to a greater or lesser degree, by increased  $\gamma$  chain expression and hence increased levels of HbF. In the  $\delta\beta$  thalassaemias,  $\gamma$  chain production is insufficient to overcome the thalassaemic phenotype and clinical disease occurs in the homozygous state or in compound heterozygotes carrying a  $\beta$  thalassaemia gene. Conditions with HbF production high enough to give good compensation are known as hereditary persistence of foetal haemoglobin and homozygotes are clinically unaffected while compound heterozygotes with  $\beta$  thalassaemia are very mild.

#### 4.2.1 Recessive mutations causing $\beta$ thalassaemia

Mutations of the  $\beta$  globin gene generally involve single base substitutions or small insertions or deletions within the gene or its immediate flanking sequences and affect almost every known stage of gene expression. These mutations have been catalogued by Huisman et al. and the updated listing is accessible electronically through the Globin Gene Server Website: <http://globin.cse.psu.edu>.



*Transcriptional mutations.* Mutations have been described in each of the CACCC, CCAAT and TATA boxes in the promoter. These result in moderate reductions in transcription to give  $\beta^+$  thalassaemias. Of the mutations in this category, the C→T mutation at position -101 to the  $\beta$  globin gene is particularly mild, such that heterozygotes are "silent" with borderline reduced/normal red cell indices.

*RNA processing mutations.* The primary transcript of the  $\beta$  globin gene is modified at the Cap site by the addition of a 7mG residue and at the 3' end by an extension of 20-50 adenine residues, the polyA tail. A mutation of the cap site (+1 A→C) produces a mild form of  $\beta$  thalassaemia while numerous mutations of the polyA addition signal result in a more severe  $\beta^+$  thalassaemia in which normally modified mRNA is reduced to ~10% of normal.

The primary transcript also has to be spliced to remove the intervening sequences. A GT dinucleotide marks the 5' (donor) end of the intron and an AG dinucleotide is found at the 3' (acceptor) end of the intron. These dinucleotides are obligatory for splicing and 20 mutations involving these invariant bases result in  $\beta^0$  thalassaemia, with either retention of the intron in the RNA or activation of cryptic splice sites in the intron or the adjacent exon. The resulting abnormal RNAs may be unstable or, if translated, result in premature stop codons and hence are non-functional.

The invariant splice nucleotides are part of a larger splice consensus sequence that involves adjacent exon and intron sequences. Point mutations of these consensus sequences reduce the rate of normal splicing resulting in  $\beta^+$  thalassaemias. In many cases they activate cryptic splice sites in the exon or intron surrounding the normal splice site producing detectable but non-functional alternative RNAs.

A further group of splicing mutations involve nucleotide replacements within introns that create a new splice site that is used preferentially to the normal site. Depending on the strength of the new site these may result in either  $\beta^+$  or  $\beta^0$  thalassaemias. Included in this group was the first characterised  $\beta$  thalassaemia allele, IVS-1 110 G→A, the commonest allele in the eastern Mediterranean. New splice sites may also be created by nucleotide substitutions in exons. These may also produce an amino acid change to produce an abnormal globin chain, albeit one produced in reduced amounts. The commonest example of this is the G→A change in codon 26 that produces the  $\beta^E$  chain of HbE. HbE is very common in south-east Asia and was known to be associated with a mild thalassaemic phenotype. Only when the associated splicing deficit was identified was this explained.

*Mutations affecting translation.* Six mutations of the translation initiation codon, AUG, have been described, all resulting in  $\beta^0$  thalassaemia. Downstream AUG

sequences are not a good match to the consensus sequence for translation initiation and would result in a shifted reading frame and a premature stop codon. Nonsense mutations in which an amino acid codon is mutated to one of the chain termination codons are a frequent cause of  $\beta^0$  thalassaemia. These include  $\beta^{039}$ , CAG→TAG, the commonest allele in the western Mediterranean. Although transcription of this gene appears normal, little or no mRNA is found in the cytoplasm as a result of nonsense mediated decay. A check point in the transport of the RNA from the nucleus detects the premature stop codon and destroys the RNA to prevent translation of truncated products that could act as a dominant negative protein.

The deletion or insertion of one or a few nucleotides (other than multiples of three) into the coding region of the  $\beta$  globin gene causes a shift in the reading frame. These frame shift mutations will allow continued reading of an altered amino acid sequence until a new termination codon is reached, producing a truncated product that is non-functional. Many of these RNAs are also likely to be subject to nonsense mediated decay.

*Deletions affecting only the  $\beta$  globin gene.* Compared to the  $\alpha$  thalassaemias, deletions that remove the  $\beta$  globin gene are comparatively rare. The 619bp deletion that removes the 3' end of the  $\beta$  gene is a relatively common allele in India and Pakistan. Most of the remaining dozen or so are restricted to single families and are necessarily  $\beta^0$  thalassaemias. They vary in length from 290bp to over 65kb. Those that remove the promoter of the  $\beta$  gene show unusually high levels of HbA<sub>2</sub> (> 6%) in heterozygotes, presumably because there is no competition between the  $\delta$  and  $\beta$  genes for transcription factors or access to the LCR.

*LCR deletions.* Large deletions that affect the entire  $\beta$  globin gene cluster are known as ( $\epsilon\gamma\gamma\delta\beta$ )<sup>0</sup> thalassaemias. They include rare case in which the entire cluster and its upstream LCR is deleted and those that delete the LCR but spare some or all of the structural globin genes but which are unexpressed. Heterozygotes for both have the phenotype of  $\beta$  thalassaemia trait but with normal levels of HbA<sub>2</sub>. They often have a severe clinical course as neonates that resolves in early childhood.

#### 4.2.2 Dominant mutations causing $\beta$ thalassaemia

*Highly unstable  $\beta$  chain variants.* Amino acid substitutions that alter the structure of the  $\beta$  chain to make it too unstable to form  $\alpha\beta$  dimers or tetramers may manifest as a dominant form of  $\beta$  thalassaemia, producing a clinical phenotype of thalassaemia intermedia. This may also occur as a result of the insertion or deletion of one or more entire codons. The unstable chain precipitates in the cell

leaving an  $\alpha$  chain excess which is also unstable. While a normal red cell is capable of dealing with the  $\alpha$  globin chain excess in a  $\beta$  thalassaemia heterozygote, having to remove the unstable  $\beta$  chain as well as the excess  $\alpha$  chains overwhelms the cell's proteolytic capacity. Undigested chains precipitate as inclusion bodies and cause severe membrane damage leading to ineffective erythropoiesis and haemolysis. In many cases the variant is so unstable that no abnormal globin can be detected in the red cells.

## 5. $\delta\beta$ thalassaemias and HPFH

### 5.1 Phenotypes

The  $\delta\beta$  thalassaemias and HPFHs are heterogeneous conditions characterised by a loss of both  $\delta$  and  $\beta$  gene expression accompanied by increased  $\gamma$  gene output. Heterozygotes for  $\delta\beta$  thalassaemia have reduced red cell indices similar to those of  $\beta$  thalassaemia but their haemoglobin pattern has normal levels of HbA<sub>2</sub> with raised HbF levels of 5-15%. Homozygotes have anaemia with a moderately severe clinical course (thalassaemia intermedia) and 100% HbF. Compound heterozygotes with  $\beta$  thalassaemia are also severely affected. HPFH was originally distinguished from  $\delta\beta$  thalassaemia by the higher level of HbF in heterozygotes (15-30%) and the benign clinical state of homozygotes (who also have 100% HbF) and compound heterozygotes with  $\beta$  thalassaemia. Homozygotes have high normal Hb levels (15-18 g/dL) presumably because of the altered oxygen affinity of the foetal haemoglobin. Globin chain synthesis studies show that  $\gamma$  chain output does not fully compensate for the lack of  $\beta$  chains with  $\alpha/\gamma$  ratios of  $\sim 2$ . As more subtypes of these conditions became recognised, the distinctions between the two became blurred and when the molecular basis for them was established it became clear that the two were closely related.

### 5.2 Molecular basis of the $\delta\beta$ thalassaemias and HPFHs

The  $\delta\beta$  thalassaemias are almost always due to deletions that remove both the  $\delta$  and  $\beta$  globin genes. These deletions are partially compensated by an increased expression of the  $\gamma$  genes and a raised level of HbF. In  $\text{G}\gamma^{\text{A}}\gamma (\delta\beta)^0$  thalassaemia the HbF contains both  $\text{G}\gamma$  and  $\text{A}\gamma$  chains whereas in  $\text{G}\gamma(\text{A}\gamma\delta\beta)^0$  thalassaemia the deletion includes the  $\text{A}\gamma$  gene and the resulting HbF contains only  $\text{G}\gamma$  chains. The sizes of the deletions vary from 9 to 100 kb and most cases have only been described from one or a few families. The Mediterranean  $\text{G}\gamma^{\text{A}}\gamma (\delta\beta)^0$  thalassaemia is more widespread and may account for up to 10% of all  $\beta$  thalassaemias in some localised regions. Crossovers between misaligned  $\delta$  and  $\beta$  genes result in a hybrid  $\delta\beta$  globin gene that results in Hb Lepore that may also be common locally.

Deletions of the  $\delta$  and  $\beta$  genes are also the molecular basis for many cases of HPFH. In general this condition is rare, as is Hb Kenya in which a hybrid  $\gamma\beta$  chain is produced. The HPFHs have higher levels of compensatory HbF production than the  $\delta\beta$  thalassaemias but a comparison of the size and the positions of the deletions has not revealed the reason some deletions produce more  $\gamma$  chains than others. Suggestions have included loss of specific regulatory regions, competition between  $\delta\gamma^A\gamma$  and  $\delta\beta$  genes and newly apposed enhancer elements. However, the  $\beta$  globin gene cluster and its surrounding sequences contain several elements with activator or repressor functions; the number and position of those elements that remain after the deletion may determine the phenotypic outcome.

Other HPFH conditions are due not to deletions but to point mutations in either the  $\delta\gamma$  or  $\delta\gamma^A\gamma$  gene promoters, resulting in HbF levels that vary between 3-30% in heterozygotes, depending on the mutation. They alter the binding of one or more transcription factors but whether they remove repressive functions or induce activator binding (or both) remains to be determined. In these non-deletion HPFH conditions, the HbF contains mainly  $\delta\gamma$  or  $\delta\gamma^A\gamma$  chains depending on which promoter is mutated. There is  $\beta$  chain production from the gene in *cis* to the mutation but the level is lower than normal and is reciprocal to the amount of  $\gamma$  chain production so that total output of the affected chromosome approximates normal.

## 6. Rare forms of $\alpha$ and $\beta$ thalassaemia

There are rare forms of  $\alpha$  thalassaemia with unusual causes. A deletion which removes the  $\alpha 1$  gene and extends into the downstream widely expressed gene which is transcribed from the opposite strand. Transcription of this newly apposed gene extends back into the  $\alpha 2$  gene and silences it (9). A single nucleotide polymorphism in the region between the  $\zeta$  and  $\alpha$  genes creates a new GATA-1 binding site that acts as a promoter, outcompeting the  $\alpha$  gene promoters and producing HbH disease in homozygotes (10).

The output of the  $\alpha$  globin genes may also be down-regulated by mutations in the X-linked ATR-X gene. Inherited mutations in this gene produce a syndromal condition characterised by mental retardation and  $\alpha$  thalassaemia (11) whereas acquired mutations are associated with myelodysplastic syndromes which also have features of  $\alpha$  thalassaemia (12).

Cases with a  $\beta$  thalassaemia phenotype unlinked to the  $\beta$  globin gene have also been described. Mutations in the erythroid transcription factor GATA-1 may have reduced  $\beta$  globin production in association with thrombocytopenia (13). Trichothiodystrophy due to mutations in the XPD gene which produces a subunit of the general transcription factor TF11H also have reduced levels of  $\beta$  globin synthesis (14).

## References

1. Weatherall DJ, Hemoglobinopathies worldwide: Present and future. *Curr Mol Med* 2008; 8: 592-599.
2. de Laat W, Grosveld F. Spatial organization of gene expression: The active chromatin hub. *Chromosome Res* 2003; 11: 447-459.
3. Higgs DR, Vickers MA, Wilkie AO et al. A review of the molecular genetics of the human alpha-globin gene cluster. *Blood* 1989; 73: 1081-1104.
4. Chui DH, Fucharoen S, Chan V. Hemoglobin H disease: Not necessarily a benign disorder. *Blood* 2003; 101: 791-800.
5. Vichinsky EP, MacKlin EA, Wayne JS et al. Changes in the epidemiology of thalassemia in North America: A new minority disease. *Pediatrics* 2005; 116: e818-825.
6. Chui DH, Wayne JS. Hydrops fetalis caused by alpha-thalassemia: An emerging health care problem. *Blood* 1998; 91: 2213-2222.
7. Higgs DR, Wood WG. Long-range regulation of alpha globin gene expression during erythropoiesis. *Curr Opin Hematol* 2008; 15: 176-183.
8. Thein SL. Genetic modifiers of the beta-haemoglobinopathies. *Br J Haematol* 2008; 141: 357-366.
9. Tufarelli C, Stanley JA, Garrick D et al. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet* 2003; 34: 157-165.
10. De Gobbi M, Viprakasit V, Hughes JR et al. A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* 2006; 312: 1215-1217.
11. Gibbons RJ, Higgs DR. Molecular-clinical spectrum of the ATR-X syndrome. *Am J Med Genet* 2000; 97: 204-212.
12. Steensma DP, Higgs DR, Fisher CA et al. Acquired somatic ATRX mutations in myelodysplastic syndrome associated with alpha thalassemia (ATMDS) convey a more severe hematologic phenotype than germline ATRX mutations. *Blood* 2004; 103: 2019-2026.
13. Yu C, Niakan KK, Matsushita M et al. X-linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. *Blood* 2002; 100: 2040-2045.
14. Viprakasit V, Gibbons RJ, Broughton BC et al. Mutations in the general transcription factor TFIIF result in beta-thalassaemia in individuals with trichothiodystrophy. *Hum Mol Genet* 2001; 10: 2797-2802.

## Multiple Choice Questionnaire

To find the correct answer, go to <http://www.esh.org/iron-handbook2009answers.htm>

1. **A patient has the following features: Hb 12.5 g/dL, MCV 68 fl, MCH 21.5pg, HbA<sub>2</sub> 2.8%, HbF 0.7%. What is the most likely diagnosis?**
  - a) Heterozygous  $\delta\beta$  thalassaemia

- b) Heterozygous  $\beta$  thalassaemia .....
- c) Heterozygous  $\alpha$  thalassaemia .....
- d) Normal .....

**2. What specific phenotypic feature is associated with heterozygous deletions of the  $\beta$  globin gene promoter?**

- a) MCH < 20pg .....
- b) HbF > 5% .....
- c) HbA<sub>2</sub> > 6% .....
- d) Hb < 10 g/dL .....

**3. Which of the following is an example of a thalassaemic haemoglobinopathy?**

- a) HbE .....
- b) HbC .....
- c) HbS .....
- d) HbD .....

**4.  $\alpha$  thalassaemia ameliorates  $\beta$  thalassaemia major through which mechanism?**

- a) Increased HbF production .....
- b) Reduced globin chain imbalance .....
- c) Increased MCV .....
- d) Decreased MCH .....

**5. Which of the following best distinguishes homozygous  $\delta\beta$  thalassaemia from homozygous HPFH?**

- a) Haemoglobin analysis .....
- b)  $G\gamma/A\gamma$  ratio .....
- c) Haemoglobin level .....
- d) % HbF .....