SUMMARY

This study on the project AT-4-074 entitled "Aspects of Human Haemoglobins and Haemoglobinopathies in the Arabian Peninsula - Studies at Genetic & Molecular Level" (The main and the two extension periods i.e. 14.12.1402 (1st October 1982) to 30.11.1413 (21 May 1993).

Screening studies

During the course of the study, different areas of Saudi Arabia were screened to determine the prevalence of anaemias, and the frequency of abnormal haemoglobin genes, the thalassaemia genes, and red cell enzymopathies genes. A total of 35,926 samples were collected and studied. The samples were collected during field studies and trips made to these areas by the principal investigator and members of research team. Special analyses other than routine haematological parameters, red cell indices and preparation of slides were carried out in Riyadh.

Anaemia in Saudi Arabia

The results show that anaemia is a frequently identified abnormality in all regions of Saudi Arabia, though the prevalence varies in different parts of the country, with the lowest prevalence in the central province and the highest in the south-western province. Microcytic-hypochromic anaemia is the most frequent findings in most of the regions, though normochromic-normocytic anaemia is also of common occurrence. In all regions screened both genetic and acquired anaemias were identified at an appreciable frequency.

Frequency of sickle cell gene and possible modulators of sickle cell gene expression in Saudi Arabia

Among the haemoglobin structural disorders the most frequent genetic abnormality was the sickle cell gene (Hb S) in all areas of Saudi Arabia, but at a variable frequency. Studies conducted on Hb S heterozygotes and homozygotes from different regions of Saudi Arabia revealed significant differences in the clinical presentation and haematological findings. In the majority of the Hb S homozygotes, mild disease was identified in the eastern province and a severe disease in the rest of the country. Possible factors involved in the amelioration of the sickle cell disease including Hb F level, associated thalassaemias, associated G-6-PD deficiency, the Gγ/Aγ ratio, the DNA polymorphisms in the β-globin gene cluster, the β-globin gene haplotypes and the locus control region were investigated. A significant ameliorating effect was observed in sickle cell disease patients with Saudi-Indian haplotype, Xmn I polymorphic site 5' to ε gene, Hpa I site 3' to β-globin gene and a high Gγ/ Aγ ratio. The number of (AT) repeats in the locus control region differed in the benign and severe disease, though none of the factors seemed to influence the Hb F level significantly.

Frequency and presentation of α- and β-thalassaemia in Saudi Arabia

Both α- and β-thalassaemias were identified in all the regions though at a variable frequency. The molecular studies revealed that the majority of the α-thalassaemias was caused by one or two α-gene deletions resulting in heterozygous (-α/αα) homozygous (-α/-α) α thalassaemia-2 states. Most of the deletion was caused by rightward deletion of a 3.7 kb fragment, though cases of leftward deletion of 4.2 kb deletion were identified, both as homozygous and heterozygous cases. Double heterozygous rightward-leftward deletion was a rare finding. No cases of hydrops fetalis were identified, but cases of Hb H disease

were encountered. The clinical, haematological and biochemical findings in the α-thalassaemia carrier were not significantly different from that of the normal non-thalassaemics except a state of mild hypochromic-microcytic anaemia which was a frequent finding in the former.

The β-thalassaemia major and trait cases were identified in all regions. The formers presented with severe clinical, haematological and biochemical abnormalities and were blood transfusion-dependent for their survival, while the β-thalassaemia trait presented with a variable degree of anaemia with a few clinical complications. The molecular basis of β-thalassaemia was investigated using Amplification Refractory Mutation System (ARMS) and dot blot analysis. The main mutations were IVS 1 nt 110, IVS 2 nt 1, C39, IVS 1 nt 5, IVS 13' end (-25), Fr 8-9, Cap+1 (A→C) and C6. Studies were conducted to determine genotype-phenotype correlations in patients with β-thalassaemia major.

Triple α-gene and β-globin gene duplication in Saudis

Cases of triple α-gene ($\alpha\alpha\alpha^{anti\ 3.7}$) were identified alone and in association with sickle cell gene. In addition, β-globin gene duplication (ββ/β) possibly resulting from unequal crossing-over, was reported for the first time in several cases from different regions of the country.

Red Cell Enzymopathies in Saudi Arabia

Among the red cell enzymopathies, the most frequently encountered enzymopathy was glucose-6-phosphate dehydrogenase (G-6-PD) deficiency, which occurred in all areas of Saudi Arabia. The enzyme deficiency was mainly caused by G-6-PD-Mediterranean,

though G-6-PD-A$^-$ was identified at a low frequency. In all regions the frequency of G-6-PD deficiency was significantly higher in the females compared to the frequency expected using Hardy-Weinberg equilibrium. No major clinical abnormality was found associated with the G-6-PD deficiency in the steady state.

Abnormal gene-gene interactions in Saudis

Gene-gene interactions occur due to co-existence in the same locality of a high prevalence of two or more abnormal genes. The interactions between genes frequently result in modification of the clinical presentation and haematological findings in the diseased states. The most frequent interactions found during this study were those resulting from sickle cell gene with α-thalassaemia, β-thalassaemia and G-6-PD deficiency, where cases with 3 or more abnormal genes were frequent, particularly in the western region.

Molecular therapy of sickle cell disease

Studies were conducted to investigate the effect of treatment of sickle cell disease and thalassaemia patients with piracetam (children) and hydroxyurea with or without erythropoietin (adult). The results showed significant improvement in the overall clinical severity of the disease by the use of piracetam, as well as hydroxyurea. All patients, particularly the non-responders showed significant improvement when erythropoietin was incorporated in the protocol. Laboratory variables showed definite improvement compatible with the clinical conditions. Studies on children who were given piracetam for treatment of sickle cell disease, where a double blind study was conducted using placebo, showed no significant effect on the haematological or biochemical parameters.

Awareness, care, control and prevention of blood genetic disorders in Saudi Arabia

To ensure proper management and care of patients suffering from sickle cell disease, a Sickle Cell Anaemia and Allied Syndrome (SAS) study group, comprised of clinicians from different medical disciplines, was formed to act as a "Referral and Consulting Body". A booklet for management of sickle cell disease and the thalassaemias was prepared and circulated to the different regions of the country. A National Working Group was formulated with members from different parts of the country and meetings of the group were held every year for the last 4 years, with the aims to improve the overall management of the patients and to follow up steps initiated towards control and prevention. These stratified service for control and management give rise to a comprehensive network comprising a primary health care, referral regional hospitals and the National Referral & Consulting Unit. The overall achievements of this study was recognised by the World Health Organisation and the Centre at Riyadh (Medical Biochemistry Department) was designated as the W.H.O. Collaborating Centre for Haemoglobinopathies, Thalassaemias and Enzymopathies, thus linking these structures to the W.H.O. regional and international network.

Publications, presentations at symposia and other relevant activities

Manuscripts were prepared at all stages during the course of this study and results were published in refereed National and International journals. In addition, papers prepared from the results of this study were presented at National and International symposia and conferences. On the other hand, a monograph on the management of sickle cell disease and thalassaemias was prepared and distributed free-of-charge to clinicians all over the Kingdom. Five booklets were prepared in Arabic for improving the general public

awareness about blood genetic disorders. Two workshops, for laboratory staff and clinicians, were held to train health team on clinical management and techniques of diagnosis of blood genetic disorders. Relevant symposia were held notably "Applied Medical Research: Past, Present and Future" held in 1983 and 1985, Symposium on DNA Technology and Genetic Engineering (1991), Symposium on Blood Genetic Disorders (1992) and the Medical Genetics in the Setting of the Middle Eastern Populations (1993). The projects study and results formed the nucleus of all these specialised scientific symposia.

Project main outcome and future prospectives

During the project AT-4-074 a major consideration was directed towards epidemiological studies, investigations of disease pathophysiology, clinical presentation and study of the molecular nature of the genetic defects, gene-gene interactions, interaction with environmental factors, and investigation of ways to decrease the suffering of these patients by better management and care strategies. Programmes for control and prevention were devised and implemented. The conception of the SAS group played a role in the success of these activities. In addition, the designation of our unit as W.H.O. Collaborating Centre for Haemoglobinopathies, Thalassaemias and Enzymopathies played a significant role in the Referral & Consulting Scheme which have enabled towards a better health care delivery, improved awareness, improved management, vaccination and prophylaxis strategies and genetic counselling. This has also helped in standardization and continuation of such activities at the national level. Symposia were held to update the knowledge and recent developments in these fields. Workshops have been carried out to train more

personnel to conduct efficient studies on blood genetic disorders, not only in Saudi Arabia but in other Middle Eastern countries.

It is contemplated that our efforts will enable effective health care delivery and will contribute positively to overcome the burden created by blood and other genetic problems.

A similar pattern of investigation for determining the magnitude of other genetic problems is complementary to these investigations. The model that has been arrived at during this study may be followed for development of facilities and strategies for the study of other genetic problems prevalent in the Saudi population.

Multifactorial and polygenic disorders are of similar chronic nature to that of red cells genetic disorder and require very similar approach for control and prevention as well as care provision to the effected individuals and share the significance of finding preclinical markers. Molecular epidemiology and telecommunications may be of relevance and would result in unveiling of possible diagnostic probes and enable monitoring of disease states, respectively.

To this end we recommend a comprehensive approach for investigation, care, management, control and prevention of genetic disorders making use of the results, the facilities and the expertise that have been developed as a result of this project.

Conclusion

In conclusion, this study covers comprehensive investigation of the haemoglobino-pathies, thalassaemias and enzymopathies in the population of Saudi Arabia. These include epidemiology, investigations of pathophysiology, molecular defects and study of the natural history of the red cell genetic disorders. A latter phase incorporated clinical trials

and the developments of comprehensive network for control, care and prevention scheme. The studies showed that genetic blood disorders are a major health problem in several regions of Saudi Arabia and play a significant role in influencing the health status in the setting up of the Saudi population. Steps towards control and prevention are, therefore, essential and must be adopted as early as possible in order to decrease the burden on the Health Authorities, the families and the community.

CHAPTER 1

GENERAL INTRODUCTION: BASIC CONCEPTS

1.0 GENERAL INTRODUCTION - BASIC CONCEPTS

Blood genetic disorders include a number of different inherited abnormalities which affect the stability and integrity of the red cells. The major defects are those that affect the haemoglobin structure or synthesis and those that affect the red cell metabolism. The haemoglobinopathies and thalassaemias are due to abnormalities in the structure or rate of biosynthesis of the globin chains of haemoglobin resulting in excessive haemolysis (Weatherall & Clegg, 1981; Williams et al, 1983), while enzymopathies effect the rate of metabolic pathways necessary for the stability and integrity of the red cells (Beutler, 1990). The haemoglobinopathies, thalassaemias and enzymopathies together constitute the major causes of inherited anaemias in human populations, particularly in areas that have a past or present history of malaria endemicity (Livingstone 1983; Luzzatto et al, 1969).

The genetic disorders of haemoglobin particularly sickle cell anaemia, α- and β-thalassaemias and the complex disorders that result from interactions of the different types of haemoglobin disorders constitute the most frequently encountered genetic diseases in man. According to a World Health Organization report, by the year 2000 approximately 7% of the world population will be carriers for one or more of these abnormal genes (W.H.O. Report, 1987).

Significant advancement had been made in the study of the molecular genetics of the haemoglobin disorders even before the advent of the techniques of molecular biology. The structure and amino acid sequence of haemoglobin molecule was worked out by application of the techniques of protein chemistry and by the study of abnormal haemoglobin variants the defect in several of these variants were identified. However, with

the advent of the techniques of molecular biology tremendous progress has been made in the understanding of the molecular genetics of the haemoglobin disorders. In this section, a brief overview of the structure of haemoglobin is provided as an introduction to the genetic defects which cause the haemoglobin disorders.

1.1 The Normal Haemoglobins

The red cells, with the major function of transporting oxygen from the lungs to the tissues, are especially equipped with the unique protein haemoglobin. The red cells contain different types of haemoglobin. The haemoglobins constitute almost 90% of the solute of the erythrocytes and function to transport the inhaled oxygen from the lungs to the different tissues, where it binds CO_2 and transports it to the lungs to be exhaled. Each haemoglobin is a tetramer of two types of globin chains arranged in a tetrahedral manner. The haemoglobins at different stages of development include Hb Gower I ($\zeta_2\varepsilon_2$), Hb Gower II ($\alpha_2\varepsilon_2$) Hb Portland I ($\zeta_2\gamma_2$) fetal haemoglobin (Hb F ($\alpha_2\gamma_2$) Hb A_2 ($\alpha_2\delta_2$) and Hb A ($\alpha_2\beta_2$). These haemoglobins differ in their globin chains, where each haemoglobin is made up of two pairs of unlike chains (Table 1.1). The globin chains include the ∘ and ζ-globin chains of embryonic haemoglobins γ-globin chain of Hb F β-globin chain of Hb A, ζ-globin chains of Hb A2 and α-chains of Hb F, A and A_2. Except the embryonic haemoglobins Gower I and Portland I, all others are made up of 2 α- and 2 non-α-globin chains, arranged forming a tetrahedron. Each globin has a haem binding pocket on the surface which holds the haem and constitutes the oxygen binding site. The haemoglobin molecules are ellipsoidal with four haem molecules on the surface where they function by combining reversibly with oxygen (Lehmann &

Table 1.1: Heamoglobin Types and their Genes

Globin Genes Chromosomes 16	11	Gene Product (Globin)	Tetramers in RBCs	Name of Hb	Present During	Conc. in adults %
	β	β-Chain	$\alpha_2\beta_2$	Hb A	Adult life	96 - 97
α	δ	α δ-Chian	$\alpha_2\delta_2$	Hb A_2	Adult life	2.5 – 3.5
	γ	γ-Chain	$\alpha_2\gamma_2$	Hb F	Fetal life	< 1.0
	ε	ε-Chain	$\alpha_2\varepsilon_2$	Hb-Gower II	Embryonic stage	0
	ε	ε-Chain	$\zeta_2\,\Sigma_2$	Hb-Gower I	Embryonic	0
ζ			$\zeta_2\,\gamma_2$	Hb Portland	stage	0
	γ	γ-Chain				

1.1.1 Chromosomal Organization of the globin genes

The globin genes are organized in two clusters; the β-globin gene cluster is located on the short arm of chromosome 11 and the α-globin gene cluster is located on the short arm of chromosome 16 (Figure 1.1).

1.1.1.1 The α-globin cluster

The α-globin gene cluster is composed of the α_2- and α_1-globin genes arranged in the order 5'-ζ-ψζ-ψα-α_2-α_1-3'. The pseudogenes ψζ and ψα-genes have a significant sequence homology to the functional ζ and α-genes, respectively, however, they are not expressed as the control of gene expression is lost (Proudfoot & Maniatis, 1980).

The α_2 and α_1 globin genes both code for the α-globin chains of fetal and adult haemoglobins. Recently a 0 globin gene was identified located 3' to the α_1-globin gene and is responsible for the synthesis of a messenger RNA identified in fetal erythroid tissue cells. However, no globin chains have so far been identified, which is coded by this gene (Maniatis et al, 1981.

1.1.1.2 The β-globin gene cluster

The β-globin gene cluster is located on the short arm of chromosome 11 and is composed of the embryonic ε-gene, two fetal yyy -genes and adult δ and β globin genes arranged in the order 5'-ε-Gγ-Aγ-ψβ-δ-β-3'. A pseudogene homologous to the β-globin genes (i.e ψβ) also exist in the cluster but is not expressed (Efstratiadis et al, 1980; Maniatis et al, 1981).

1.1.2 Developmental Alterations

During the different stages of development several changes take place in the

Figure 1.1: The α- and β-globin gene clusters

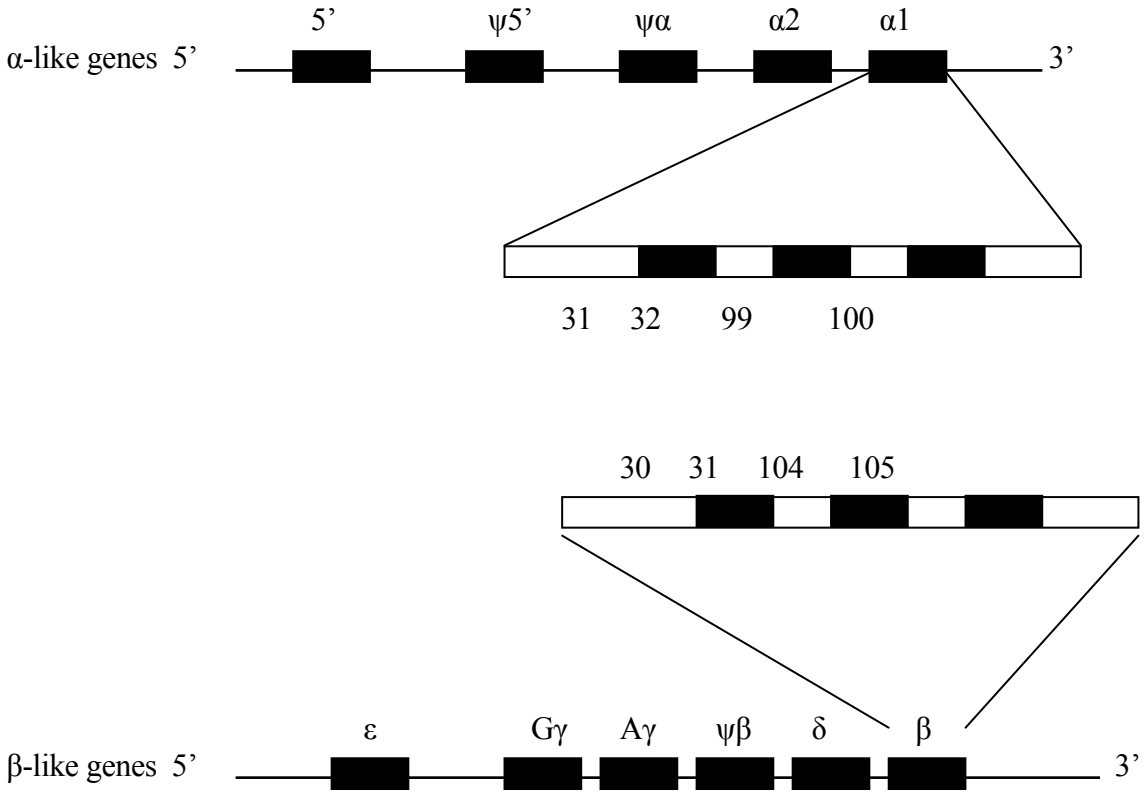
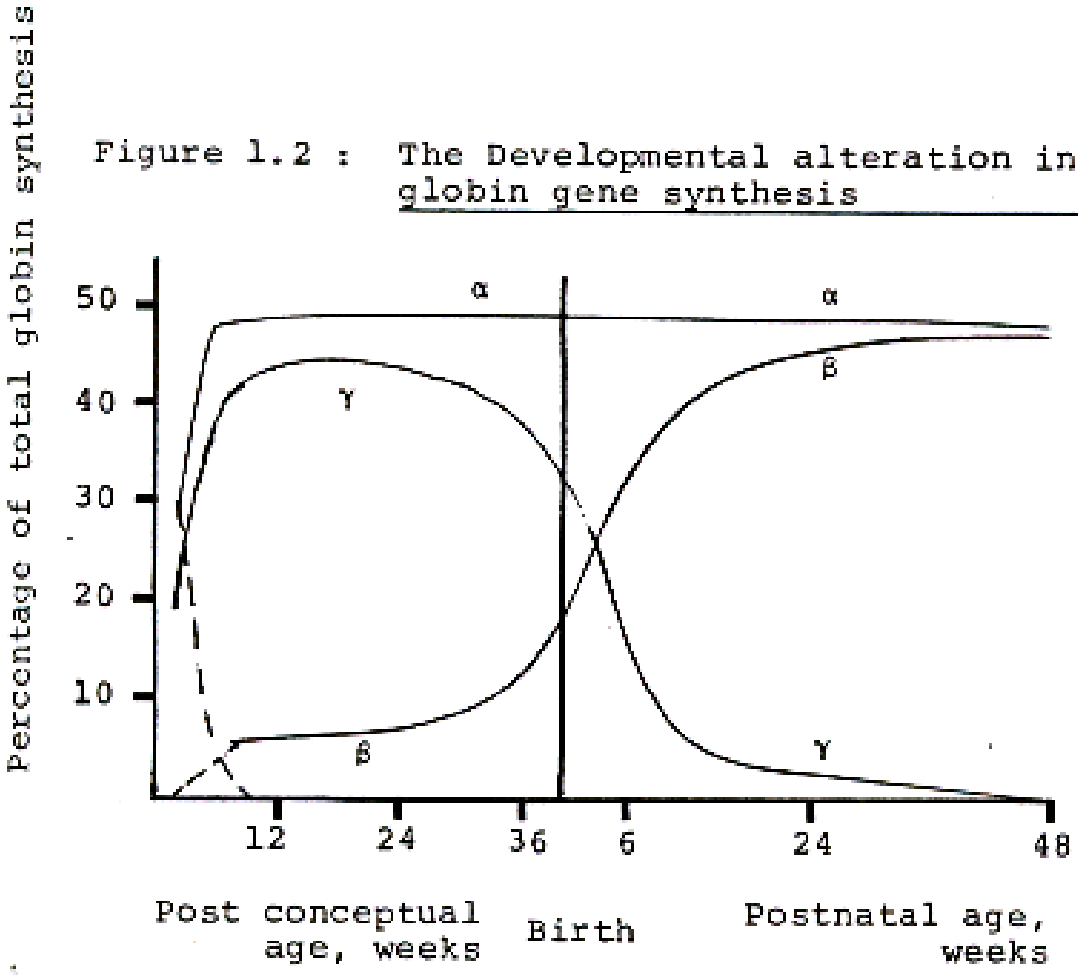


Figure 1.2 : The Developmental alteration in globin gene synthesis



expression of the globin genes and in the major site of erythropoiesis. These changes are summarized in Figure 1.2. During the early embryonic period ζ and ε-genes are expressed and form haemoglobin Gower I, Portland I and Gower 2. The erythropoiesis occurs in the yolk sac until the age of about five weeks of gestation. At this stage, the site of erythropoiesis shifts to the liver and is accompanied at about the same time with decrease in the synthesis of ζ and ε-chains and an increase in the synthesis of α- and γ globin chains of fetal haemoglobin ($\alpha_2\gamma_2$) As fetal development proceeds, the bone marrow progressively becomes the major site of erythropoiesis and at the time of birth it is the major erythropoietic organ. During the last several weeks before birth, a shift also takes place from the expression of γ-globin gene to β-globin genes of adult haemoglobin, Hb A ($\alpha_2\beta_2$) birth almost 75% of the haemoglobin is Hb F and the rest is Hb A. Within the first year after birth the γ-gene expression is gradually reduced while the globin gene expression increases gradually and at about the age of one year the Hb F is reduced to almost 1% and Hb A constitutes almost 97% of the total haemoglobin. After birth the γ-globin gene expression also occurs and Hb A_2 reaches the adult value of 2.5 - 3.5% by one year of age.

1.1.3 Structural Features of the globin genes

Each of the globin genes is made up of three exons (i.e. the coding regions) and two Introns (i.e. the intervening sequences, IVS, the non coding region) as shown in Figure 1.3 (Maniatis et al, 1981; Liebhabar et al, 1980; Lawn et al, 1980) about 100-1300 base pairs and is the initiation site for transcription by specific DNA-binding proteins. During transcription both exons and introns are transcribed for the synthesis of the primary RNA transcript, however, splicing of this RNA takes place to produce the mature mRNA which

is made up of only the exon region sequences. Prior to RNA splicing, RNA processing takes place which involves methylation of the guanine at 5' end (i.e. CAP) and the 3' end to add the poly A tail. The cap plays an important role in subsequent steps of RNA processing and during translation, while the poly A tail is necessary for the stability of the mRNA. The mature mRNA leaves the nucleus and in the cytoplasm on the ribosomes initiates the synthesis of the globin chains. The steps involved in globin gene expression are presented in Figure 1.4.

1.1.4 Control of globin gene expression

The globin genes are tissue specific and are expressed only in the erythroid cells. In addition, the expression of the different genes varies during ontogeny. This control of gene expression is achieved through specific regulatory DNA elements which bind specific protein factors. The protein - DNA interactions are not fully understood, however, with the advent of the techniques of molecular biology, a major focus of the current research is towards understanding of these control mechanisms.

The DNA regulatory elements are (a) the promoter elements and (b) the enhancer elements (Dierks et al, 1983; Salditt-Georgieff & Darnell, 1983; Proudfoot and Brownlee, 1976; Darnell, 1982).

1.1.4.1 The Promoter elements of the globin genes

The promoter elements of a gene are those DNA sequences that are essential for accurate and efficient initiation of transcription (Figure 1.5).

i) The β-globin gene promoters:

Upstream from the transcription site of the globin genes are the short highly

conserved sequences which form the promoter elements. The most proximal of these is

Figure 1.3: Globin Gene Structure

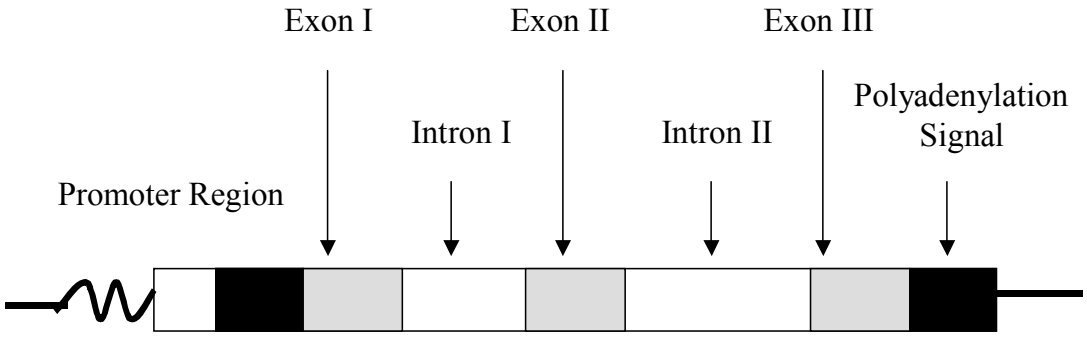


Figure 1.4: β-globin gene expression

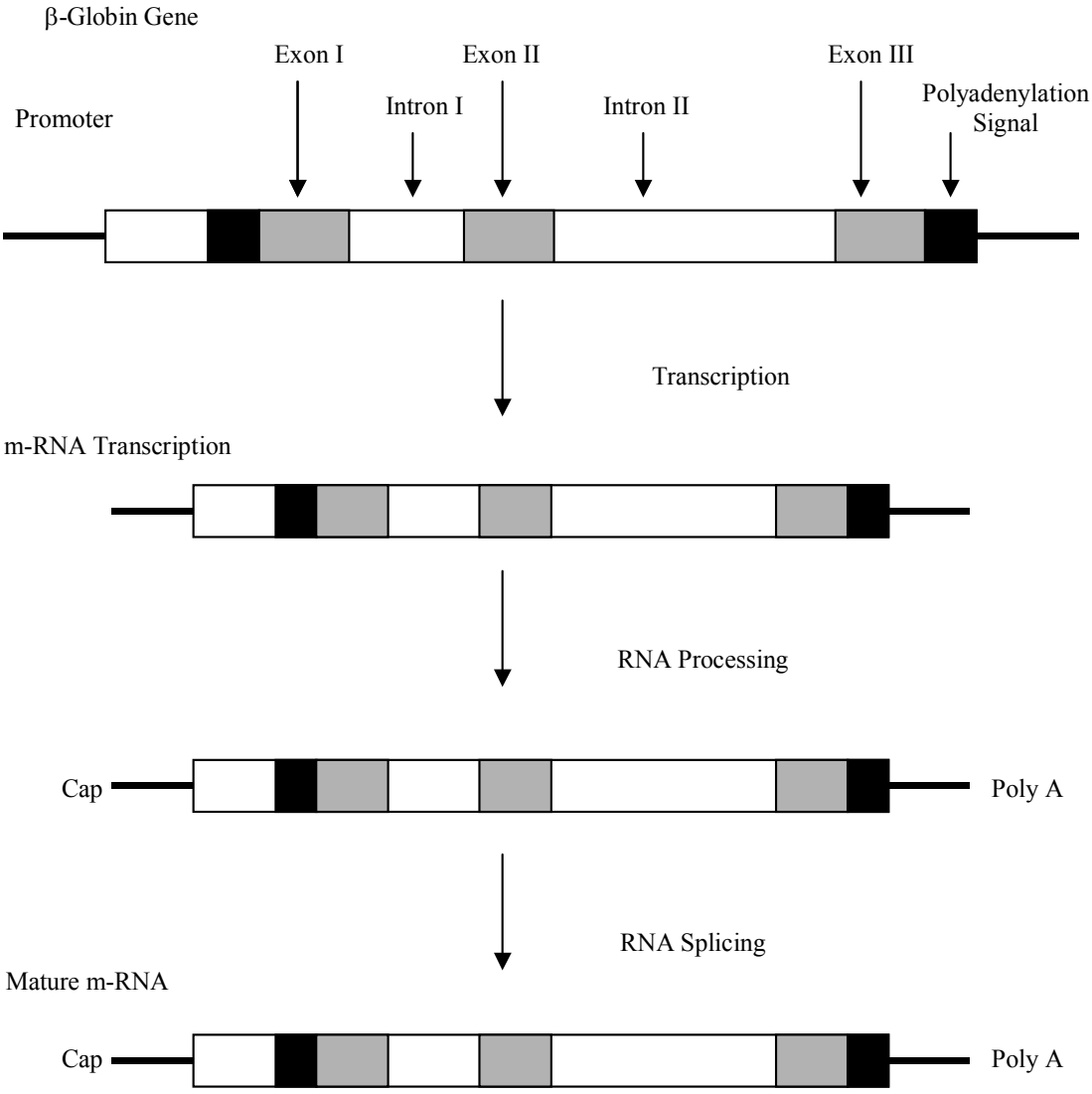
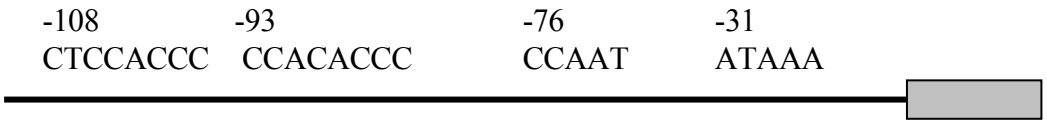
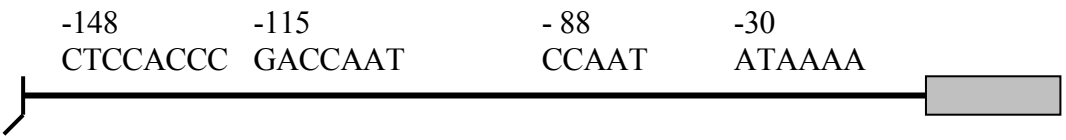


Figure 1.5: β- and γ-Globin Gene Promoters

A. β-Globin Gene Promoter



B. Aγ-Globin Gene Promoter



about -31 bp upstream from the RNA initiations site and is an AT rich region known as the TATA box. This promoter element functions to ensure accurate initiation of the mRNA transcript. The second promoter element is -76 bp upstream from the CAP site and is known as the 'CCAAT' box. This region is involved in the regulation of transcription. The third promoter elements located – 93 upstream from the cap site are 'CACCC' box. (The CACCC box is duplicated in β-globin gene). The 'CCAAT' and the 'CACCC' boxes referred to as the `upstream promoter elements', primarily serve to increase the rate of transcription. Mutations in these promoter elements result in decreased β-globin gene expression and hence in β-thalassaemias. Thus single molecular substitutions within the β-globin gene promoters sequences CCAAT and TATA boxes, or changes in the relative distance between the different promoter sequences or deletion of the promoter sequences, all result in marked decrease or complete absence of the expression of the β-globin genes. In addition, deletions of the duplicated promoter sequences (i.e. CACCC), completely abolishes promoter function, while single nucleotide substitutions can cause a substantial decrease in the amount of correctly initiated β-globin mRNA.

ii) <u>The δ-globin gene promoter</u>

The δ-globin gene promoter regions consists of the TATA box, -31 bp upstream from the cap site and the CCAAC box, -68 bp upstream from the cap site. No other promoter sequences are found in the δ-globin gene region. Alterations in the δ-globin gene promoter region have been considered as the cause of low expression of the δ-globin gene and hence a low Hb A_2 level. Such a δ-globin gene is said to be `thalassaemic'. The

major difference in the δ-globin gene promoter region compared to the β-globin gene promoter region are as follows: (i) the CAAT box is different (i.e. CCAAC instead of CCAAT), (ii) the distance between this box and the TATA box is shorter, (iii) the conserved sequences between -86 and -108 nucleotides upstream from the CAP site are lacking. Thus these defects in the promoter sequences are believed to be a major cause of hypo-function of the δ-globin gene. In addition, the δ-globin gene mRNA is found to be less stable than the β-mRNA and this further results in decreased synthesis of the δ-globin chains.

iii) The γ-globin gene promoter region:

The γ-globin gene promoter region contains the same three sequence elements as the one's present in the β-globin gene promoter i.e. 30 bp upstream in the TATA box, -88 bp upstream is the CAAT box and -148 bp upstream is the 'CACC' box. However, the 'CAAT' box is duplicated and the second CAAT box is located -115 bp upstream from the CAP site, between the 'CACC' and the first 'CAAT' box. In addition, unlike the β-globin gene upstream promoter, the CACC promoter is not duplicated in the γ-globin gene. The promoters and the nearly upstream sequences are shown to play an important role in the developmental regulation of the γ-globin gene expression. Several patients with non-deletion hereditary persistence of fetal haemoglobin have been identified and were found to have single base substitutions within and upstream from the Gγ or the Aγ

globin gene promoters. It is likely that the DNA binding proteins interact with these sequences to mediate the fetal developmental stage specific expression.

iv) The ε-globin gene promoter region:

The ε-globin gene promoter has the TATA box -28 bp upstream from the cap site, CAAT box -83 bp upstream and the CACC box -115 bp upstream. Like the γ -globin gene the CACC box is not duplicated.

Thus, within the β-like globin gene promoters there are several individual variations when compared to the β--globin gene which is the most expressed of all the globin genes during adult life. Promoter sequences are believed to play a role in the developmental regulation of the globin gene (Stamatoyannopoulos and Nienhuis 1981; Reviewed by Nienhuis et al, 1984; Weatherall & Clegg, 1981).

1.1.4.2 The enhancer elements of the globin genes

The globin gene promoter requires an 'enhancer' for its functions. These enhancers consist of short segments of DNA, that act independent of orientation or distance from the gene, to ensure accurate and efficient initiation of transcription and are, therefore, involved in gene regulations. In some cases the enhancer elements contain the same or similar sequence motifs as the promoter elements and these sequences interact with common protein factors.

The 'enhancer' elements differ from the promoter elements (i) that they are able to exert affects on genes over distances of upto several thousands base pairs, (ii) they can

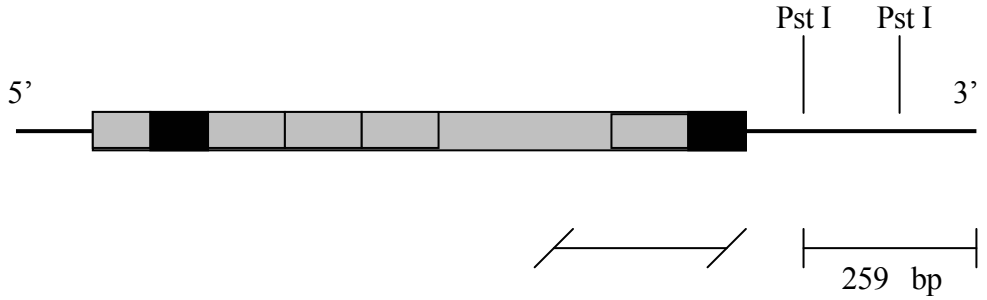
function in either orientation relative to the promoter, (iii) they are position independent and may be present 5' to, within or 3' to the gene, and (iv) some enhancers exhibit tissue or developmental stage specificity. Loss of an essential enhancer could result in loss of gene expression, as is believed to be the case in γδ-deletion chromosome with an intact β-globin gene. An enhancer was found located 55 bp upstream to β-globin gene, another 259 bp downstream to the β-globin gene regions and 744 bp region 3' to the A γ-globin gene (Figure 1.6). An intragenic β-globin gene enhancer is also identified involving the 3rd exon of the β-globin gene and djacent DNA. Both the 3'-enhancer sequence and the intragenic β-globin gene enhancer are important in developmental stage specific gene expression.

An erythroid specific protein factor, called the nuclear factor erythroid I (NFE-I) has been identified which binds to four places in the β-globin enhancer to a repeated consensus sequences. It also binds to several positions in the human β-globin promoter and in the intragenic enhancer. It is suggested that the NFE-1 and its DNA recognition sequence play an important role in regulation of β-globin gene expression.

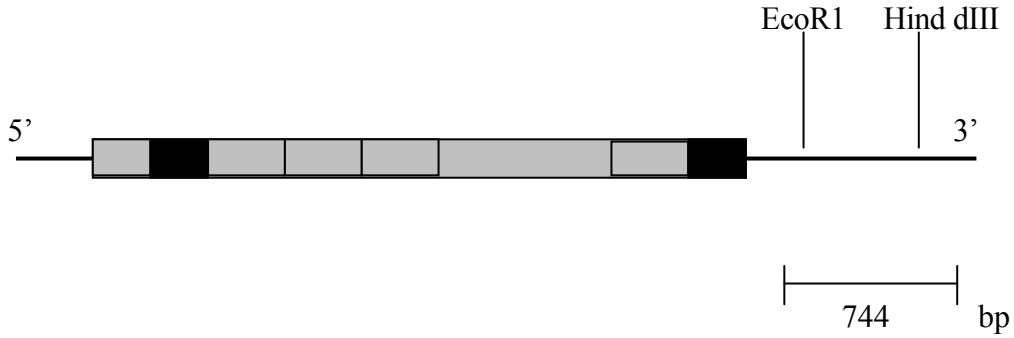
An enhancer has also been revealed 3' to the A γ-globin gene and has been shown to play an important role in increasing the A γ-globin gene expression. Other important regulatory sites have been identified 5' to within and 3' to, a number of the globin genes and it is expected that information regarding the interaction of these sequence elements with nuclear protein factors will clarify the picture of the tissue-specific and developmental stage specific regulation of globin gene expressions.

Figure 1.6: Human Globin Gene Enhancers

A. β-Globin Gene Enhancers



B. Aγ-Globin Gene Enhancers



1.1.4.3 Distant regulation of the β-globin gene cluster

Distant sequences to both the β-globin gene and the α-globin gene cluster on chromosomes 11 and 16, respectively, play an important role in control of globin gene expression. Five major erythroid-specific DNase I hypersensitive sites have been identified. Four sites located 5' to the ε-globin gene within a 20 Kb DNA fragment and one downstream to the β-globin genes cluster constitute the β-globin "locus control region" (β -LCR). These sites are developmentally stable, erythroid specific and are present at all stages of development.

These sites are located 6.1, 11.0, 14.5 and 17.5 Kb upstream from the ε-globin gene transcriptional and start site and about 21.8 Kb downstream from the β-globin gene poly adenylation signal (Figure 1.7). The consensus nomenclature now designate the most upstream 5' hypersensitive site a 5' HS4 and the others 5' HS3, 5' HS2, 5' HS1 and 3' HS1, respectively. The β-LCR was previously called the locus activation region (LAR) or dominant control region (DCR).

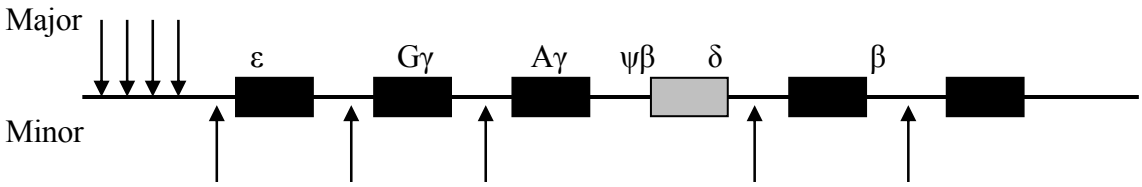
There is sufficient experimental evidence to suggest that the β -LCR in required for both erythroid and developmental specific chromatin structure. This evidence has been obtained using both transgenic mice and tissue cell cultures. Deletion mutation in the β-LCR result in inactivation of the entire β-like gene chromosomal locus despite normal genes and intact flanking regulatory sequences. In addition, the locus becomes resistant to DNase I digestion in both erythroid and non-erythroid cells.

The most important sites in the β-LCR are the tandemly arranged 5' HS2 and 5'HS3. Their activity is conserved in fragment of a few hundred base pairs which allow

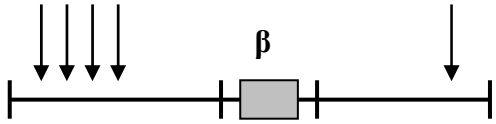
high level of position and orientation independent expression. The 5' HS3 fragment

Figure 1.7: Distant Regulation of the β-Globin Gene

A. DNAse Hypertensive Sites



B. β-Globin "mini-locus



contains NF-E2 & GATA-1 binding sites and GGTGG motives. The 5' HS2 fragment contains similar sequences plus a tandem repeat for activating protein (API) binding sites. A recent study has shown that an (AT)xN12(AT), repeat motive in this region has a similar structural and nuclear protein binding patterns to that of a silencer region of β-globin gene.

The exact mechanisms by which the β-globin gene LCR influences the β-globin gene expression is not understood. However, two distant functions have been clearly defined: (β- LCR organizes the entire β-globin locus into an open chromatic domain sensitive to DNase I, (ii) β-LCR severe as a powerful transcriptional enhancer for the whole set of β-like globin genes.

Thus the β-LCR could be a site of attachment to the nuclear matrix, or binding sites for topo-isomerase II or related enzymes or could be an origin of DNA replication or an erythroid specific enhancer of replication (Stamatoyannopoulos & Nienhuis, 1991).

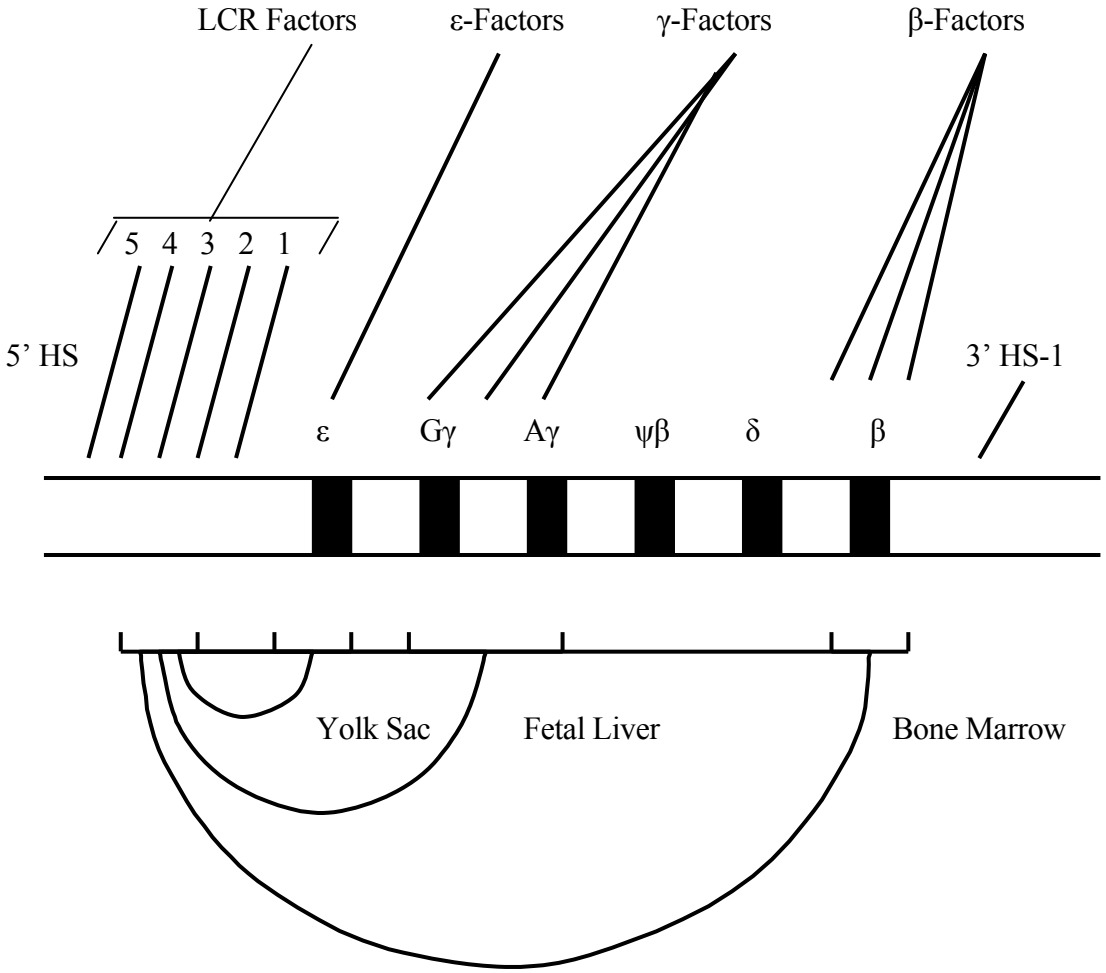
The model proposed for the stage specific regulation of the genes of the β-globin locus during the different stages of development is shown in Figure 1.8.

1.1.4.4 The Alpha-globin gene locus control region (α-LCR).

A remote tissue specific, positive regulatory element has been identified and appears to be of major importance in determining the human α-globin gene expression. This region is located about 40 Kb upstream of the ζ2 gene and so far has been found to be similar to the globin gene expression. β-LCR in its position, structure and function.

The α- and β-LCR probably provide target sequences through which they co-ordinate expression and the regulation of the α- and β-like globin genes are achieved

Figure 1.8: General Model for Haemoglobin Switching



(Reviewed by Nienhuis et al, 1984).

1.1.5 The Synthesis of haemoglobin

The different globin genes are expressed at the different stages of development. The mechanism involved in the switching-over from expression of one globin gene to the other are in the process of being clarified. The overall effect is that the α-globin- like genes synthesise mRNA for the α-globin-like chains, which leave the nucleus after splicing and other post-transcriptional modifications and induce the synthesis of the α-globin-like chains in the cytoplasm on the ribosomes. On the other hand, in the same manner the β-globin-like genes synthesise the β-globin-like chains in the cytoplasm. The haem is introduced in the globin chains and finally 2 α-globin-like and 2 β-globin-like chains are arranged together in a specific structure to give different normal haemoglobins at different stages of development (Figure 1.9).

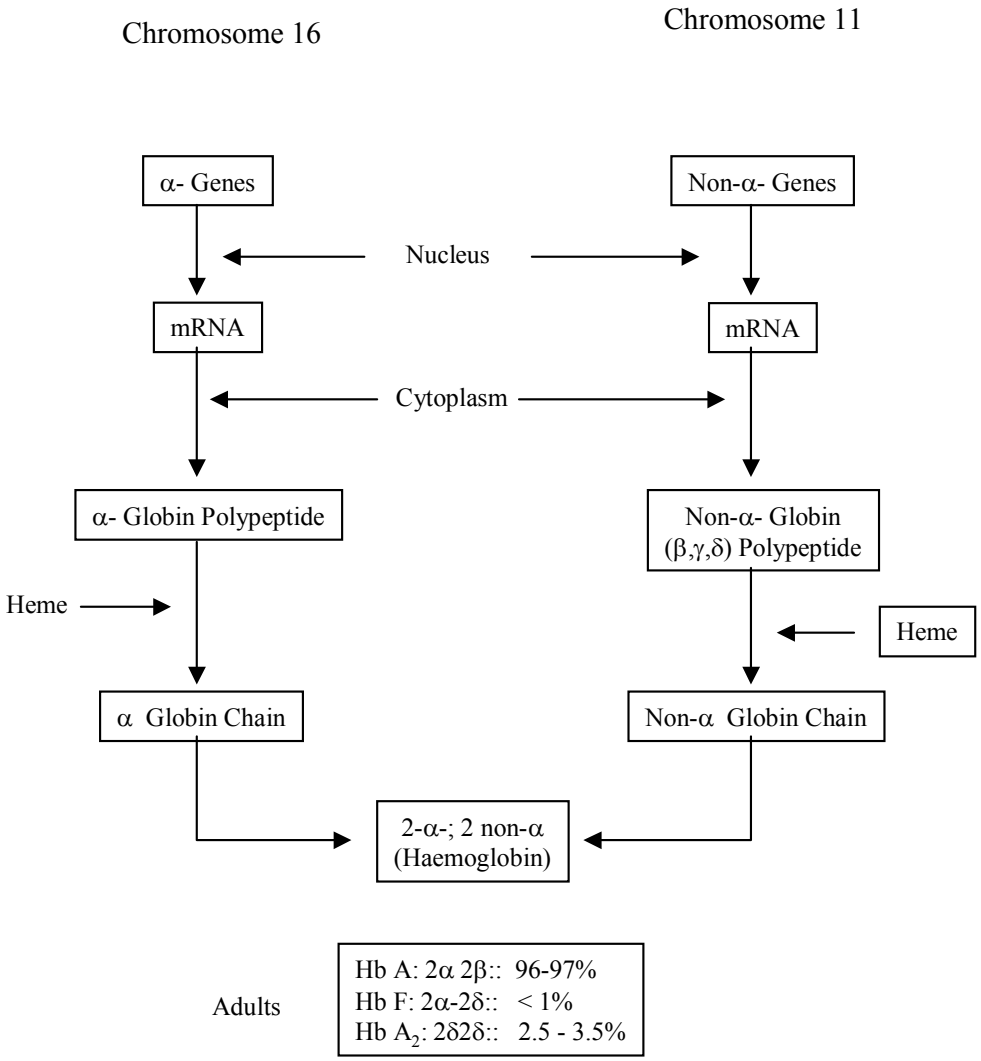
1.1.6 The normal adult haemoglobins

All types of normal human haemoglobins are tetramers of two types of globin peptides. The major adult haemoglobin is Hb A ($\alpha_2\beta_2$) which constitute about 96% of the total haemoglobin in adults. Haemoglobin A_2 ($\alpha_2\,\beta_2$) occurs at a level of 2.5 - 3.5%, while Hb F is less than 1% in the adults blood. The changes in the haemoglobin composition during development is an adaptive process designed to meet the altered oxygen transport requirements of the individual at the different stages of development. In adults variations in the level of the different haemoglobin types do not normally occur, however, they are frequent in patients with haemoglobin disorders.

Each globin chain has a specific number and sequence of amino acid which

governs its folding into a specific secondary and tertiary structure. Four of the globin

Figure 1.9: The synthesis of Haemoglobin



chains (2α and 2 non-α are compiled together in a specific three dimensional tetrahedral arrangement and are held together by non-covalent bonds. This is the quaternary structure of haemoglobin. Each globin chain has a hydrophobic cavity formed by the folding of the polypeptide chain, which holds tightly a non-protein component known as "heme". Heme is made of tetrapyrrole ring structure which are held together by methylene bridges. In the centre is a ferrous iron which forms four links with the nitrogen in the four pyrrole rings and is also linked to proximal histidine (His F8) in the globin chain by a coordinate covalent link. At the 6th co-ordinate position the ferrous iron can bind oxygen molecule and this is the oxygen binding site of the molecule (between the ferrous iron and a distal histidine, His E7). Each haemoglobin has four hemes and four oxygen binding sites and is thus referred to as an "allosteric protein" (Lehmann & Hunstman, 1974). The specific amino acid composition is determined by the specific codon sequences in the globin gene. The genes located on the chromosomes 11 and 16, synthesis the mRNA which transports the genetic information to the ribosome in the cytoplasm for the synthesise of the globin peptides. Any mutation is the gene, in the exons or in the intervening sequences results in either altered structure and function of the globin chains, altered stability, or altered rate of globin chain synthesis, thus leading to structural or biosynthetic defects of haemoglobin i.e. haemoglobinopathies (Weatherall & Clegg, 1981).

1.2 The Anaemias

The term anaemia refers to a decrease in the concentration of total haemoglobin in blood, decrease in the total number of circulating erythrocytes, or decrease in the

haematocrit level compared to the normal population. For comparison, relevant normal populations are used living in the same locality, since under different environmental conditions the normal haematological parameter values may be different e.g. a haemoglobin level normal at sea level may be relatively low at a high altitude.

Anaemias are of different types and are classified either (i) on the basis of the red cell morphology or (Figure 1.10), or (ii) on the basis of the pathophysiological mechanism involved in the pathology of the anaemia (Figure 1.11) (Chandrasoma and Taylor, 1993). On the other hand anaemias may be either genetic or acquired (Figure 1.12). The genetic anaemias result from mutation in the genes encoding haemoglobin, red cell enzymes, red cell membrane proteins or areas surrounding these genes and responsible for their expression. These disorders are inherited as `single gene disorders and follow "Mendelian Rules of Inheritance" and may be dominant or recessive. The haemoglobin disorders are generally autosomal recessive (haemoglobinopathies, thalassaemias), though a few are inherited as autosomal dominant traits. The `enzymopathies' i.e. the enzyme defects and red cell membrane defects are both autosomal and X-linked, and examples of recessive and dominant inheritance are frequent (Chandrasoma and Taylor, 1993; Williams et al, 1983; Hardisty & Weatherall, 1982; Beutler, 1990; Edwards and Bouchier, 1991; Kumar and Clark, 1990).

The acquired anaemias are a large diverse group, caused by either decreased production, or increased loss or increased destruction of blood components. The decreased production which may be due to deficiency of some essential nutrient or marrow failure or secondary to an endocrinopathy, may be reversible and is corrected

upon correction of the deficiency or it may be irreversible such as due to destruction of the bone marrow and the correction of such an anaemic state is not easy. Excessive

Figure 1.10

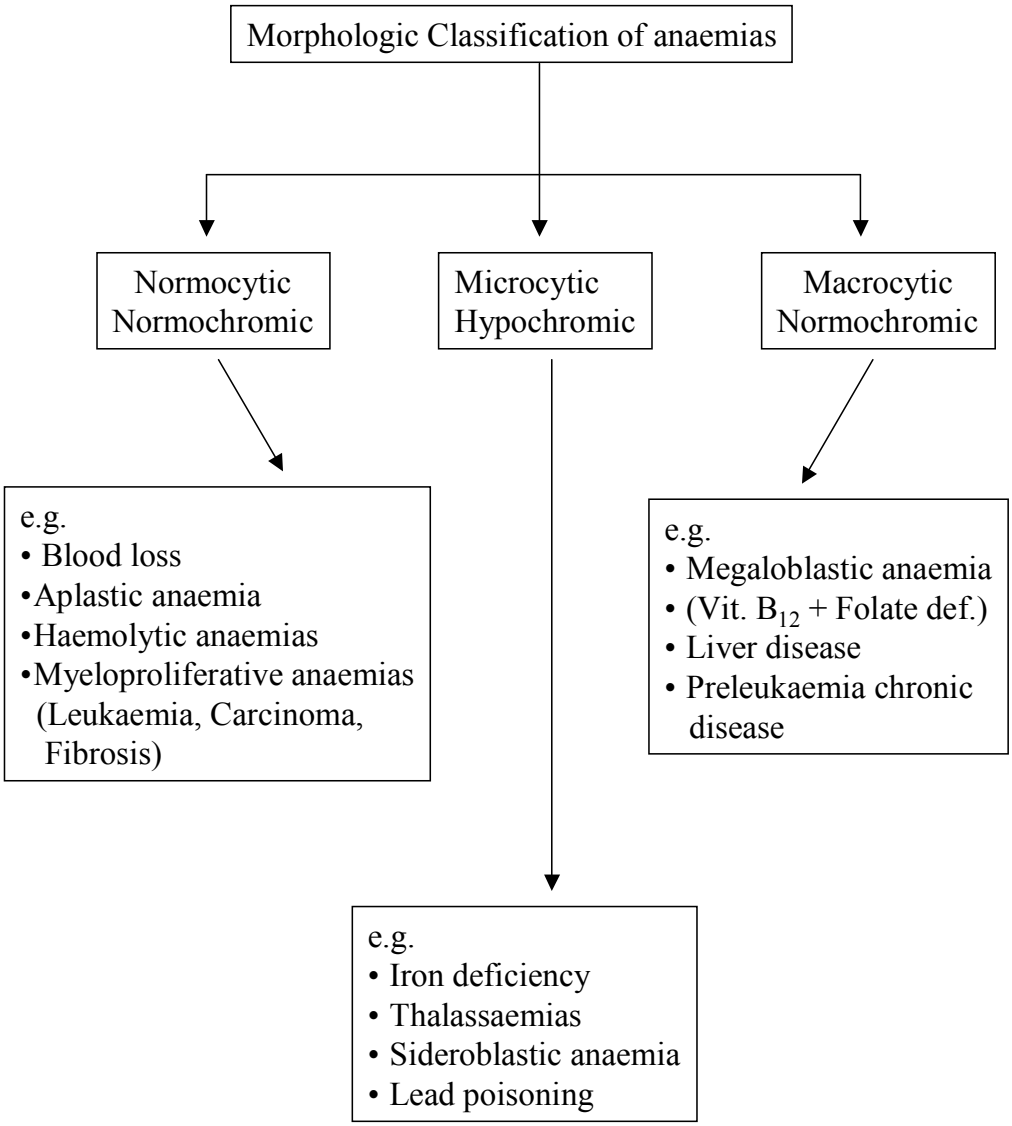

Morphologic Classification of anaemias
Normocytic Normochromic
Microcytic Hypochromic
Macrocytic Normochromic
e.g.
• Blood loss
•Aplastic anaemia
•Haemolytic anaemias
•Myeloproliferative anaemias (Leukaemia, Carcinoma, Fibrosis)
e.g.
• Megaloblastic anaemia
• (Vit. B_{12} + Folate def.)
• Liver disease
• Preleukaemia chronic disease
e.g.
• Iron deficiency
• Thalassaemias
• Sideroblastic anaemia
• Lead poisoning

Figure 1.11

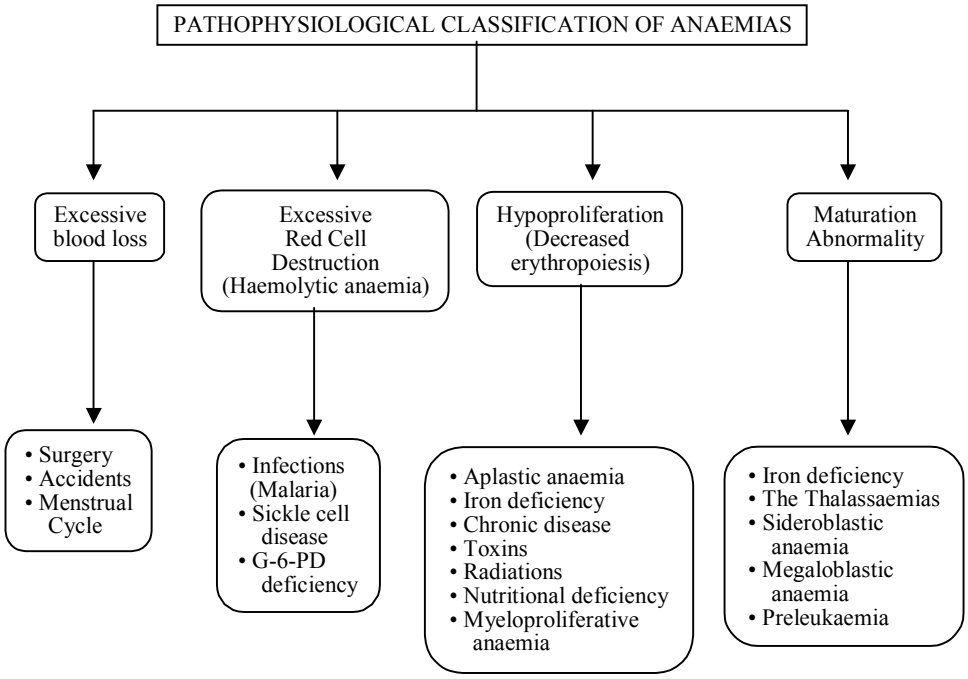
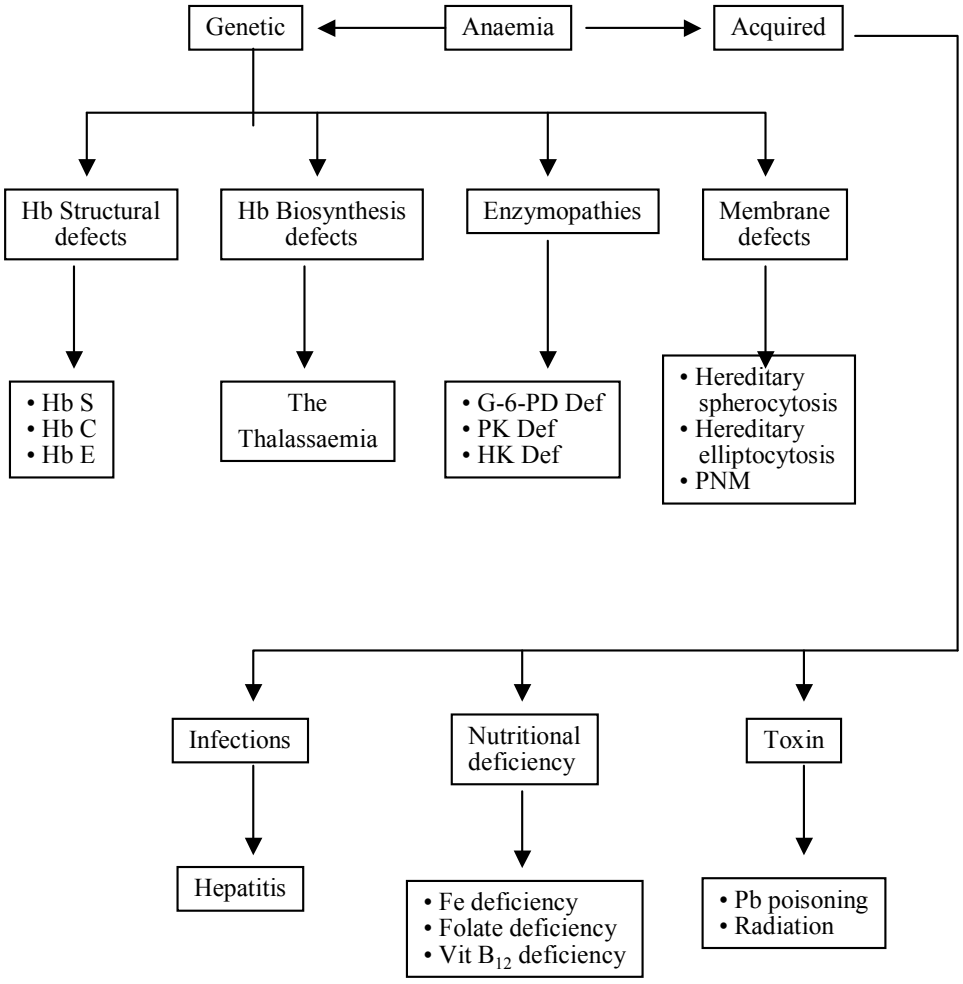

PATHOPHYSIOLOGICAL CLASSIFICATION OF ANAEMIAS
Excessive blood loss
Excessive Red Cell Destruction (Haemolytic anaemia)
Hypoproliferation (Decreased erythropoiesis)
Maturation Abnormality
• Surgery
• Accidents
• Menstrual Cycle
• Infections (Malaria)
• Sickle cell disease
• G-6-PD deficiency
• Aplastic anaemia
• Iron deficiency
• Chronic disease
• Toxins
• Radiations
• Nutritional deficiency
• Myeloproliferative anaemia
• Iron deficiency
• The Thalassaemias
• Sideroblastic anaemia
• Megaloblastic anaemia
• Preleukaemia

Figure 1.12: Genetic and Acquired anaemias



blood loss may be either acute or chronic and may result in either a mild or severe anaemia, which may even require blood transfusion for correction. The anaemia is corrected upon treatment of the basic cause of excessive blood loss. Finally, increased destruction of the red cells is a frequent cause of acquired anaemia. Various infective agents, plant or bacterial toxins, snake venom and excessive lead intoxication are some of the factors which result in excessive haemolysis resulting in acute or chronic haemolytic anaemia. Individuals with certain enzyme deficiencies (e.g. glucose-6-phosphate dehydrogenase (G-6-PD) deficiency) are more susceptible to haemolysis when exposed to certain environmental factors such as drugs, infections and *Vicia faba* compared to non deficient individuals. These are examples where both genetic and environmental factors are essential for the production of the anaemic state and in the absence of the environmental factors there is no anaemia.

1.3 Genetic Disorders of Haemoglobin

Genetic disorders of haemoglobin involve either the haem or the globin chains. Very little is know about genetic defects of haem and those that occur due to enzyme deficiencies affecting haem synthesis are rare. The defect in the globin chain structure or synthesis are referred to as "haemoglobinopathies". Types of haemoglobinopathies include: (a) those due to structural defects in the globin chains e.g. sickle cell haemoglobin; (b) those due to decreased rate of globin chain synthesis i.e. the thalassaemias; (c) those involving both structural changes and rate of synthesis i.e. the Hb E/thalassaemias; and (d) the hereditary persistence of total haemoglobin, which are characterized by persistence of Hb F synthesis during adult life.

1.3.1. The Haemoglobinopathies

The haemoglobin structural disorders result from mutations in the exons of the structural genes of the globin chains. The haemoglobin variants identified in human populations until 1990 are listed in Table 1.2. The mutation may occur in the α- β, γ, δ globin genes and results in structural changes in haemoglobin. Some of these structural alterations have no clinical consequence while others produce a haemoglobin with altered structure, functions and stability (Figure 1.13). Various degree of clinical consequences have been reported associated with these abnormal haemoglobins. These includes haemoglobins with altered oxygen affinity (increased or decreased), decrease stability and excessive haemolysis (Perutz & Lehmann, 1968; Lehmann & Huntsman, 1966; Livingstone, 1967; Ingram, 1957; 1961).

Table 1.3 lists some of the more common haemoglobin variants and the mutations that produce these variants.

1.3.1.1 The molecular genetics of haemoglobin structural disorders

The haemoglobin structural disorders result from a large number of different mutation which influence the structural genes of haemoglobin (Maniatis et al, 1980). The mutations so far reported include:

(a) Single base Substitutions

Almost 400 haemoglobin variants have been identified which result from a single base substitution and majority of these are α- and α-globin chain variants. A single base is replaced and the mutations are either a

transition (purine → purine or pyrimidine → pyrimidine) or a transversion mutation (purine → pyrimidine or vice versa). This may or

Table 1.2: Haemoglobin variants identified in human populations until 1990*

Type of Variant	Number
α-Chain	156
β-Chain	272
γ-Chain	58
δ-Chain	18
Deletions	14
Insertions	3
Deletion/Insertion	2
Fusions	9
Elongation	10
Two Point Mutations	13

* List of mutant haemoglobin

Figure 1.13: Structural variants of haemoglobin dueto mutation in the globin genes

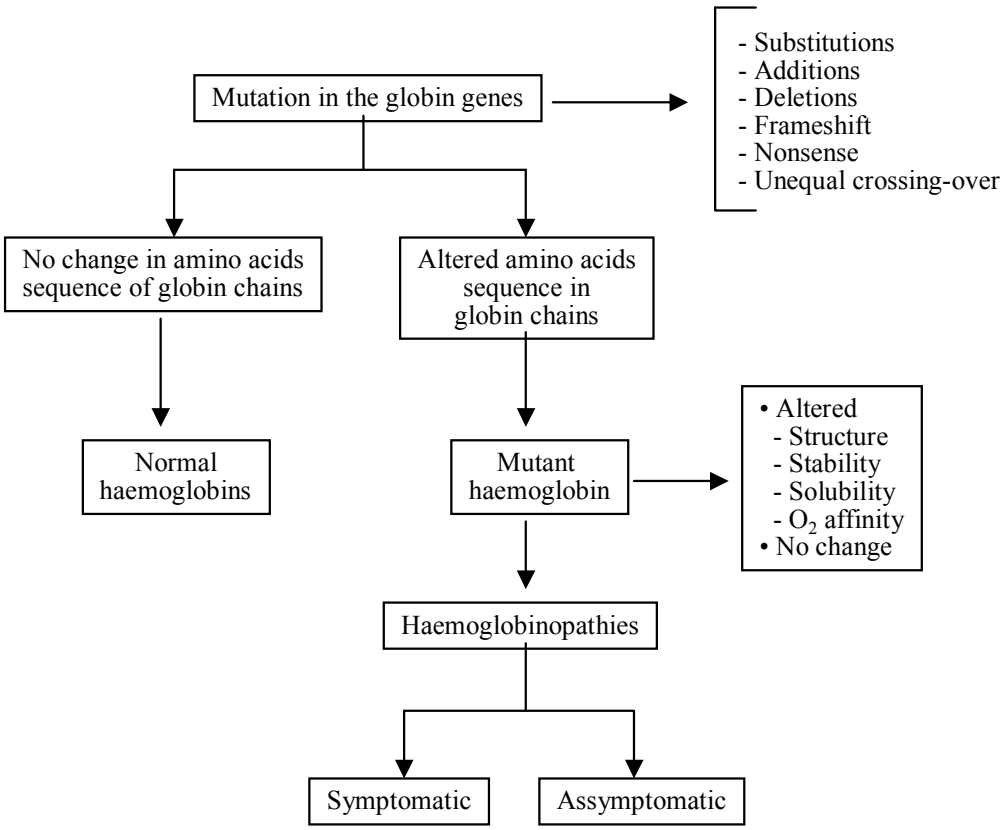


Table 1.3: Common Beta globin chain haemoglobin variants

	Mutation in gene (DNA)		Amino acid in globin chain	
	Codon No.	Codon Change	No.	Change
Hb S	6	GAG → GTG	6 (A3)	Glu → Val
Hb C	6	GAG → AAG	6 (A3)	Glu → Lys
Hb E	26	GAG → AAG	26 (B8)	Glu → Lys
Hb D	121	GAA → CAA	121 (GH4)	Glu → Lys

may not result in a change in the amino acid in the globin peptide. However, if an amino acid is changed, the alteration in the physical properties and in the functions will depend on the nature and location of the substituted amino acid.

Of the more than 400 structural variants of haemoglobin only haemoglobin S, C, D and E occur at polymorphic levels in some populations, while the others are rare and frequently not associated with any serious clinical state. They can be detected by separating the haemoglobin types on electrophoresis, if the amino acid substituted has altered the charge on the globin chains, or by isoelectric focussing if the isoelectric pH of the globins are altered. A large number of mutant haemoglobin cannot be separated from the normal haemoglobins since there is no change in the charge. Some of these are detected if they alter haemoglobin structure, function or solubility or lead to a clinical disorders.

A few variants can only be explained on the basis of two base substitutions in the same codon. These include Hb Sydney$^{(\beta 67\ Val \rightarrow Ala)}$, Hb Bristol$^{(\beta 67\ Val \rightarrow Asp)}$ and Hb Milwaukee$^{(\beta 67\ Val \rightarrow Gln)}$. It is suggested that these double substitutions result from cross-over between the respective variant allelic genes.

(b) Frame-shift Mutation Variant

Frameshift mutations are those mutations which result from addition or deletion of one or more bases (except three bases or multiples of three). In these cases the haemoglobin variant produced has a completely

altered amino acid sequence from the point of mutation i.e. the whole frame has shifted. Some of the examples of variants produced as a result of frameshift mutation are Hb Cranston, Hb Wayne and Hb Tak. In Hb Tak and Hb Cranston the mutation occurs near the C-terminal end and the globin peptide are elongated by a few amino acids by reading the normally nontranslated part of the mRNA at the 3' end. If the mutation occurs near the 5' end of the mRNA, then the protein synthesized will have such an altered amino acid sequence that it may not bind haem or to behave as haemoglobin.

(c) Deletions and Insertions

Deletion or addition of three bases or multiples of three bases results in the deletion or addition of one or more amino acid residues without any alteration in the rest of the amino acid sequence. Such deletions or additions are believed to result from unequal crossing-over at meiosis. Included in this group are several haemoglobin variants including Hb Leiden$^{(\beta 67\ Glu \rightarrow O)}$, Hb Lyon β 17-18 $^{(Lys\ Val \rightarrow O),}$ Hb Gun Hill $^{(\beta\ 92\text{-}96,\ Leu\text{-}His\text{-}Cys\text{-}Asp\text{-}Lys \rightarrow O)}$, Hb Leslie $^{(\beta 131 Glu \rightarrow O)}$ and Hb Tours $^{(\beta 87\ Thr \rightarrow O)}$.

(d) Chain Termination Mutations

Several haemoglobin variants have been described which result from a mutation in the termination codon of the globin chains of haemoglobin. Once the termination codon is changed to a codon for an amino acid, the synthesis of proteins along the mRNA continues and

additional amino acids are added, thus elongating the globin peptide. One of the first examples of this type of mutations is Hb-Constant Spring which has 172 amino acid residues instead of 141 in the α-globin chains. Hb constant spring and other α-globin chain elongation variants are thalassaemic in nature i.e. the rate of their synthesis is decreased, thus giving α-thalassaemic picture.

(e) Premature Terminations

A few haemoglobin variants have been reported which have a shorter chain length and result from conversion of a codon for an amino acid to a termination codon. Thus the polypeptide is terminated prematurely. This type of mutation is also referred to as a `nonsense mutation'. An example of premature termination is Hb McKees Rocks which is terminated at amino acid (UAU→UAA or UAG) and lacks the C-termination tyrosine and histidine in the β-globin chains. Prior to this amino acid, the whole chain has a normal amino acid sequence.

If the mutation occurs earlier on in the globin gene, the unstable haemoglobin results presenting itself as a thalassaemic phenotype.

(f) Fusions

If during crossing-over mechanisms the chromosomes are misaligned, then the resulting genes will be composed of parts of two different genes. Such mutations are referred to as fusion mutations. Several haemoglobin variants have been identified which contain normal α-chains

but the non α-chains are composed of part of δ and β or γ β-chains. These are known as the haemoglobin Lepore. The δ-amino acid sequence is present at the N-terminal end and the β sequence at the C-terminal end. One of the hybrid chromosomes has δ-β combination but has lost its normal δ and β- genes, while the other has normal δ and β-genes and in addition has β-δ hybrid which is known as Anti-Lepore. Hb Kenya has a non-α-chain which is a hybrid γ β chain of 146 amino acids. These fusion haemoglobin variants have clinical phenotype of γβ-thalassaemia.

1.3.2 The Thalassaemias

The thalassaemias are genetic disorders of haemoglobin that result from mutation in or around the globin genes and are characterized by a reduced rate of production of one or more of the globin chains of haemoglobin. Thus the thalassaemias are disorders of haemoglobin biosynthesis. Generally, the haemoglobin which is synthesized is normal with no structural defect, however, due to the imbalance between the α and non α-chains, the excessive globin chains frequently associate to form tetramers of similar chains (e.g. $\gamma_4\beta_4$). These are unstable and deposit in the red cells affecting its stability and proper functioning. The thalassaemias constitute a diverse heterogenous group of inherited defects which are generally associated with a state of hypochromic-microcytic anaemia of varying degree depending on the amount of available globin chains. The thalassaemias are classified depending upon the gene in which the mutation has taken place and whose synthesis is decreased, i.e. α-thalassaemia result from decreased synthesis of α-chains and β- thalassaemias result from decreased synthesis of β-chains. In addition α- and β-

thalassaemias are further subgrouped as α+ or β^{+}thalassaemia in which some α or β-chain synthesis occurs and α° or β°thalassaemias in which no α- or β-globin chain is synthesized (Weatherall & Clegg, 1981).

A brief outline of the classification of α- and β-thalassaemias and the mutations that produce the thalassaemias are presented in Table 1.4 to Table 1.6.

1.3.2.1 Molecular Genetics of the Thalassaemias

Deletion of the whole gene, part of the gene and different types of point mutations occurring within the intervening sequences (introns), the promoter region, the poly-adenylation signals, exons, flanking regions, regulatory regions, in remote DNA sequences and in the Locus Control Region result in decreased expression or complete absence of the globin gene expression and hence lead to the thalassaemias. Hence the molecular pathology of the thalassaemias is exceptionally diverse compared to other genetic disorders. All mutations are cis mutations i.e. they effect only the gene on the chromosome in which the mutation occurs (Weatherall & Clegg, 1981; Nienhuis et al, 1984).

(a) **The Molecular Genetics of the α-Thalassaemias**

The α-genes are a pair of linked genes (α_2 and α_1) 1ocated on the short arm of each of the chromosome 16. Each α-gene has three exons, and two introns (IVS). In the 5'-flanking region are the various regulatory boxes which are involved in the regulation of gene expression and on the 5' end are the poly A adenylation site.

Throughout development the expression of α2 gene signals, predominates that of α1 gene by 3:1.

Table 1.4: Types of the α –Thalassaemias

A: Genetic Determinants

Description	Haplotype	Heterozygous State	Homozygous State
α-Thal-1	--/	α °-thalassaemia 5-10% Hb Bart's at birth. Low MCH and MCV	Hb Bart's hydrops faetalis
α-Thal-2	-α/	α –Thalassaemia 0-2% Hb Bart's at birth. Minimal haematological change	As for heterozygous α -Thal-1
Non-deletion α-Thal	$\alpha\alpha^T$	May be similar to above but haemato-logical changes may be more severe	Hb H disease in some cases

B: Genetic Interactions

Interaction	Genotype	Disorders
α -Thal-1/ α -Thal-2	---/-α	Hb H Disease
α -Thal-1/ non-deletion α-thal.	--/α α	Hb H Disease
Non-deletion α-Thal/non-deletion α -Thal	$\alpha\alpha^T/\alpha\alpha^T$	Hb H Disease
Dysfunctional α -Thal/ α -Thal-2	-α/-α	Hb H Disease

Table 1.5: Major types of the β-thalassaemias

Type	Homozygote	Heterozygote
β°	Thal Major No β-Cain (No Hb A)	Thal minor - 5% Hb A_2
β+ Mediterranean	Thal Major - 10-20% Hb A - 50-70% Hb F	Thal minor - 5% Hb A_2
β+ (Negro	Thal Intermedia - 30-50% Hb F - 50-70% Hb A	Thal minor -5% Hb A_2

Table 1.6: Major types of mutations that produce the Thalassaemias

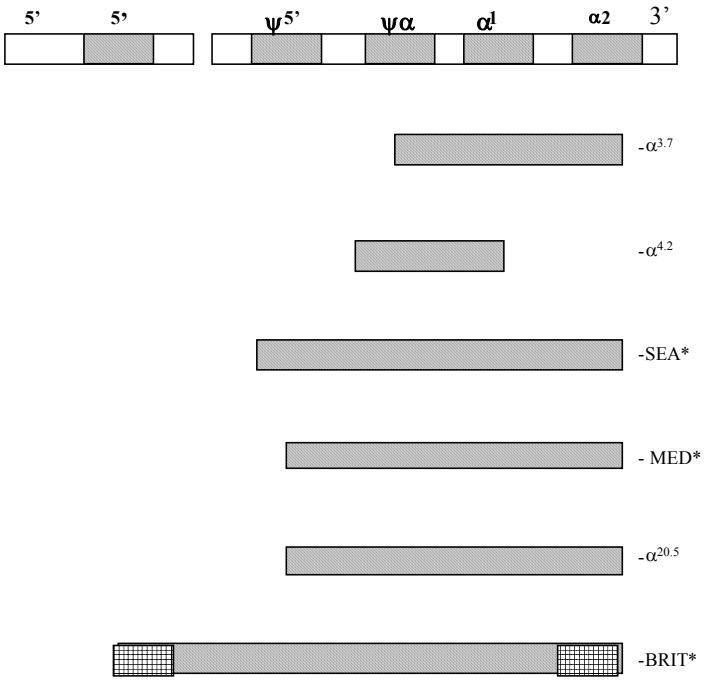
- β-Thalassaemia:
  - ♦ Deletions
    - ♦ 3.7 Kb (rightward deletion)
    - ♦ 4.2 Kb (leftward deletion)
    - ♦ α°-mutations
  - ♦ Non-deletion mutations
    - Splice junction mutations
    - Single point mutation.
    - Structurally abnormal globin
    - Polyadenylation signal mutation
    - Chain termination mutations
    - One, two or five base pair deletions (frameshift)
- β-Thalassaemias:
  - ♦ Point mutations (may produce β° or β^{+}-Thal
  - ♦ Non-functional mRNA
    - Nonsense mutations
    - Frameshift mutations
  - ♦ RNA processing mutations
    - Splice junction changes
    - Consensus changes
    - Internal IVS changes
    - Coding region substitutions affecting processing
  - ♦ Transcriptional mutations
  - ♦ Cap site mutations
  - ♦ Unstable globins
  - ♦ Premature chain terminations
  - ♦ Gene deletions
    - α°-Thal.
    - (δβ)°-Thal.
    - (γδβ)°-Thal.
    - (εγδβ)°-Thal

i) Deletion α-Thalassaemia (Figure 1.14)

The majority of the α-thalassaemias result from deletion of various segments of one or more of the α-gene chains. The commonest types of α-Thalasaemias are those resulting from deletion of one of the duplicated α -genes (-α/; α^+ Thal), Deletion of both of the gene (--/; α°-Thal) occur frequently in the population of the Mediterranean region and in South-east Asia (Weatherall & Clegg, 1981; Higgs et al, 1990).

In patients with α°-thalassaemia there is complete absence of α -chain synthesis. In the homozygous state i.e. hydrops faetalis (--/--), there is no α-chain synthesized as the total α -gene content is deleted and the condition is characterized by severe intrauterine hypoxia with still birth or death soon after birth. In this condition, the α-containing HbA&F are completely absent and there is detectable Hb Portland ($\zeta_2\gamma 2$) at birth. The combined double heterozygous state resulting from inheritance of α°-thal and α^+-thal produces a condition referred to as Hb H disease (-α/--) Hb H (β_4) is detectable during electrophoresis and the condition gives rise to severe anaemia compatible with survival into adult life. Deletion of one or two αgenes (i.e. the heterozygous α-thalassaemia 2 (-α/αα), the heterozygous α-thalassaemia 1 (--/αα) and the homozygous α-thalassaemia 2 (-α/-α) are the most frequently

Figure 1.14: Molecular Pathology of deletion forms of α-Thalassaemia



*SEA: South East Asian
*MED: Mediterranean
*BRIT: British

encountered α-thalassaemia states. The conditions are generally asymptomatic with varying degree of hypochromic-microcytic anaemias.

Some deletions involve only the α-globin gene area, while the others cover much longer portions of the α-globin gene cluster, not only the two α-genes but also the ψα-genes and the ζ-gene may be deleted.

The most frequent of these deletions are those due to deletion of 3.7 Kb fragment involving a rightward deletion. In some populations the leftward deletion i.e. deletion of 4.2 Kb fragment has been reported. This type of deletion may be produced as a result of unequal crossing-over of non-aligned chromosome, producing the deletion on one chromosome and triple α-gene ($\alpha\alpha\alpha^{anti\ 3.7}$ or $\alpha\alpha\alpha^{anti\ 4.2}$) on the other. The triple α-gene arrangement is generally rare and indicates that other mechanisms may be involved in the α-gene deletions.

The α°-thalassaemia result from deletion of different lengths of the α-globin gene cluster extending from the α-globin gene to the ζ-globin gene.

Recent studies have shown an association between α-gene deletions and mental retardation. Several patients were described who had α-thalassaemia and mental retardation with a variety of

dysmorphic features. They were shown to have large deletions of the entire α-globin complex. In some of these patients the deletion was believed to result from imbalanced chromosome translocation, while in the other cases the mechanism was not clear. It was suggested that there might be an important locus in this region which is responsible for the associated mental retardation.

Another group of 5 patients with α-thalassaemia were identified, who also had mental retardation. However, the α-genes were intact and the condition was labelled as α-thalassaemia/mental retardation non-deletion. This condition was different from the one described above.

A form of α-thalassaemia has been recognized which results from a deletion of a large (62 Kb) fragment, upstream of the α2-globin gene and the affected chromosome is referred to as $(\alpha\alpha)^{RA}$ since both the α-genes and their transcription units are intact. The possible mechanism/s involved are not clear. However, it has been suggested that the positive regulatory element of the α-gene which are necessary for their expression are deleted and hence result in inhibition of α-gene expression and hence α-thalassaemia phenotype. A similar mechanism that has been shown to produce β-thalassaemia involves deletions in the regulatory sequences which severely down regulates the β-globin

gene expression and hence produces β-thalassaemia

(ii) The non-deletion α-Thalassaemias (Figure 1.15)

Alpha-thalassaemias caused by point mutations or short deletions have been reported to result in α-thalassaemia, though at a lower frequency than the deletion type of α-thalassaemias. These include RNA processing mutations, mutations in the RNA termination codons and mutations in the exons affecting the translation of the mRNA producing unstable mRNA. In each of these situations the α -genes are intact but the mRNA is either not processed properly or translated in the normal way.

There is a family of α-thalassaemias which result from a single point mutation in the α-globin chains termination codon. As a result of this the termination codon changes to a codon for an amino acid and the α-globin chain synthesis continues until the next stop codon is reached i.e. after additional 30 aminoacid. One of the first haemoglobin, identified with an elongated α-chain is Hb Constant spring. The stop codon UAA changes to CAA, a codon for glutamine.