HAD1 8/5/05 2:11 PM Page 1

1 Haemoglobin and the genetics of haemoglobin synthesis

Haemoglobins and their structure and function

The haemoglobin molecule contained within red blood cells is essential for human life, being the means by which oxygen is transported to the tissues. Other functions include the transport of carbon dioxide (CO_2) and a buffering action (reduction of the changes in pH that would otherwise be expected when an acid or an alkali enters or is generated in a red cell). A normal haemoglobin molecule is composed of two dissimilar pairs of polypeptide chains, each of which encloses an iron-containing porphyrin designated haem (Fig. 1.1). Haemoglobin has a molecular weight of 64–64.5 kDa. Haem is essential for oxygen transport while globin serves to protect haem from oxidation, renders the molecule soluble and permits variation in oxygen affinity. The structure of the haemoglobin molecule produces an internal environment of hydrophobic radicals which protects the iron of haem from water and thus from oxidation. External radicals are hydrophilic and thus render the haemoglobin molecule soluble. Both haem and globin are subject to modification. The iron of haemoglobin is normally in the ferrous form (Fe^{2+}). Haem is able to combine reversibly with oxygen so that haemoglobin can function as an oxygen-transporting protein. Oxidation of iron to the ferric form (Fe^{3+}) is a less reversible reaction, converting haem to haematin and haemoglobin to methaemoglobin, a form of haemoglobin that cannot transport oxygen.

The haemoglobin molecule can also combine with CO_2, being responsible for about 10% of the transport of CO_2 from the tissues to the lungs; transport is by reversible carbamation of the N-terminal groups of the α chains of haemoglobin. Carbamated haemoglobin has a lower oxygen affinity than the non-carbamated form, so that binding of the CO_2 produced by the metabolic processes in tissues facilitates oxygen delivery to tissues. In addition, non-oxygenated haemoglobin can carry more CO_2 than oxygenated haemoglobin, so that unloading of oxygen to the tissues facilities the uptake and transport of CO_2. Because of its buffering action (mopping up of protons, H^+), haemoglobin also contributes to keeping CO_2 in the soluble bicarbonate form and thus transportable. The reaction $CO_2 + H_2O \rightarrow HCO_3^- + H^+$ is facilitated.

The haemoglobin molecule has a role in nitric oxide (NO) transport and metabolism. Haemoglobin is both a scavenger and an active transporter of NO. NO is produced in endothelial cells and neutrophils by the action of nitric acid synthase [1–3]. NO has a very high affinity for oxyhaemoglobin, so that blood levels are a balance between production and removal by binding to oxyhaemoglobin. NO is a potent vasodilator, this effect being limited by its binding to haemoglobin. The iron atom of a haem group of oxyhaemoglobin (preferentially the haem enclosed in the haem pocket of an α chain) binds NO. A haemoglobin molecule with NO bound to two haem groups strikingly favours the deoxy conformation, so that oxygen is readily released. NO-haemoglobin is subsequently converted to methaemoglobin with release of NO and the production of nitrate ions, which are excreted. As deoxyhaemoglobin has a much lower affinity for NO, hypoxic conditions may leave more NO free and lead to vasodilation of potential physiological benefit.

NO also causes S-nitrosylation of a conserved cysteine residue (Cys^{93}, E15) of the β globin chain of oxyhaemoglobin to form S-nitrosohaemoglobin. This occurs in the lungs. In this circumstance, the bioactivity of NO may be retained, with NO being delivered to low molecular weight thiol-containing molecules to reach target cells, such as the smooth muscle of blood vessels. Oxygenation of haemoglobin favours S-nitrosylation. Conversely, deoxygenation favours

HAD1 8/5/05 2:11 PM Page 2

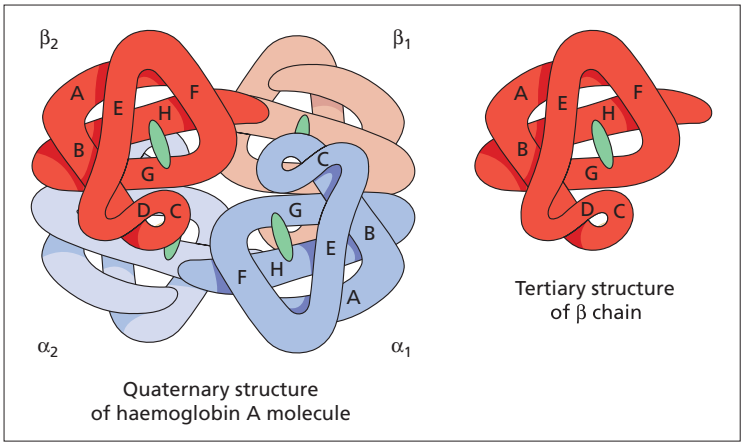

Fig. 1.1 Diagrammatic representation of the tertiary structure of a haemoglobin monomer (a β globin chain containing a haem group) and the quaternary structure of haemoglobin; upper case letters indicate homologous α helixes.

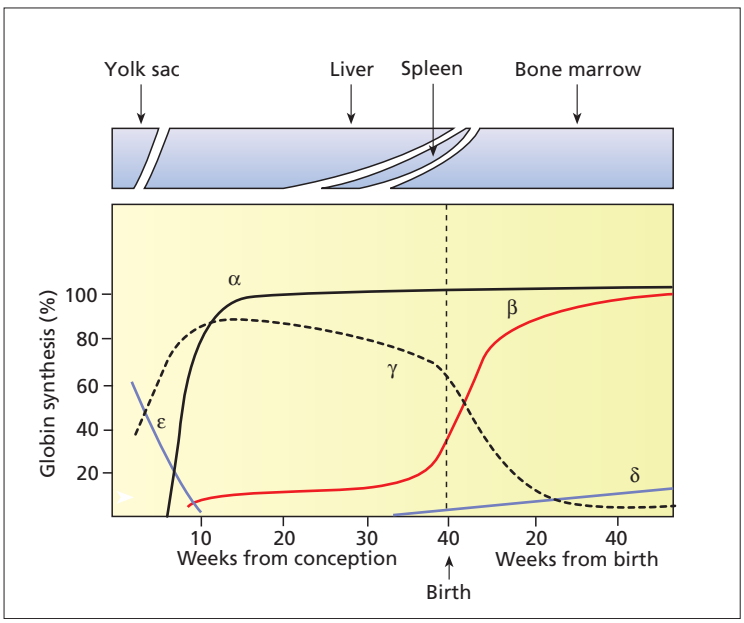

Fig. 1.2 Diagrammatic representation of the sites and rates of synthesis of different globin chains *in utero* and during infancy.

the release of NO. This may be an important physiological process, with NO being released in peripheral tissues where it can facilitate arteriolar dilation. The oxy form of S-nitrosohaemoglobin is a vasoconstrictor, whereas the deoxy form is a vasodilator. Lack of oxygen could thus again favour vasodilation.

In normal circumstances, the ability of haemoglobin to scavenge or destroy NO is reduced by the barrier to NO diffusion provided by the red cell membrane; however, in chronic haemolytic anaemia, increased free plasma haemoglobin may lead to impaired vascular responses to NO [3]; inactivation of NO by haemoglobin in the plasma may contribute to the pulmonary hypertension that can be a feature of sickle cell anaemia, and also the hypertension that has been observed with some haemoglobin-based blood substitutes.

As a result of the synthesis of different globin chains at different stages of life (Fig. 1.2), there is a difference in the type of haemoglobin present in red

HAD1 8/5/05 2:11 PM Page 3

Table 1.1 Haemoglobins normally present during adult, fetal and embryonic periods of life.

Haemoglobin species	Globin chains	Period when normally present
A	$\alpha_2\beta_2$*	Major haemoglobin in adult life
A_2	$\alpha_2\delta_2$	Minor haemoglobin in adult life; even more minor in fetal and neonatal life
F	$\alpha_2{}^{G}\gamma_2$ or $\alpha_2{}^{A}\gamma_2$	Minor haemoglobin in adult life; major haemoglobin in fetal life with a declining percentage through the neonatal period
Gower 1	$\zeta_2\varepsilon_2$	Significant haemoglobin during early intrauterine life
Gower 2	$\alpha_2\varepsilon_2$	Significant haemoglobin during early intrauterine life
Portland or Portland 1†	$\zeta_2\gamma_2$	Significant haemoglobin during early intrauterine life

*Can also be designated $\alpha^{A}{}_2\beta^{A}{}_2$ to distinguish the globin chains of haemoglobin A from those of variant haemoglobins.
†Haemoglobin Portland 2 ($\zeta_2\beta_2$) has been observed in α thalassaemia syndromes, but is unlikely to occur in significant amounts during normal development.

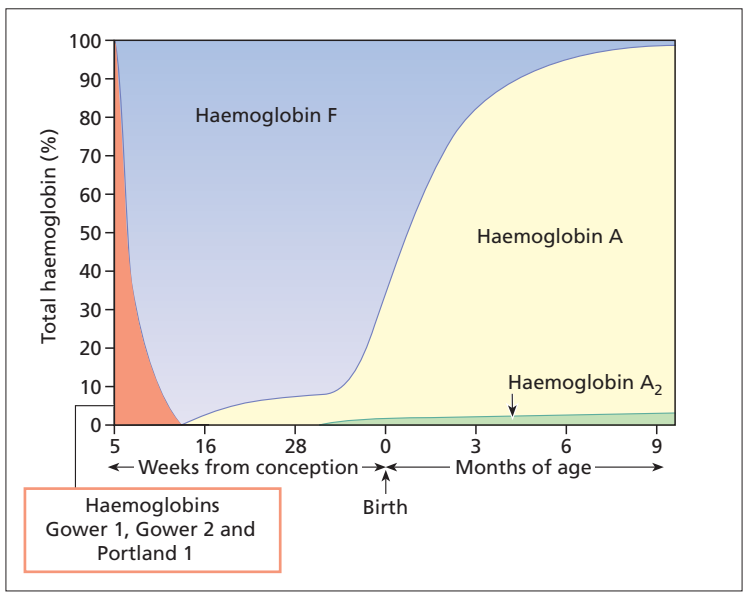

Fig. 1.3 Diagrammatic representation of the average percentages of various haemoglobins present during the embryonic and fetal periods and during infancy.

cells between adult life and the fetal and neonatal periods (Table 1.1, Fig. 1.3). In adults, 96–98% of haemoglobin is haemoglobin A (A = adult), which has two alpha (α) chains and two beta (β) chains. The name 'haemoglobin A' was given by Linus Pauling and colleagues in 1949 when they discovered that asymptomatic carriers of sickle cell disease had two different haemoglobins, which they designated haemoglobin A and haemoglobin S [4]. A minor haemoglobin, haemoglobin A_2, has two α chains and two delta (δ) chains. Its existence was first reported in 1955 by Kunkel and Wallenius [5]. A very minor haemoglobin in adults, but the major haemoglobin during fetal life and the early neonatal period, is haemoglobin F or fetal haemoglobin, which has two α chains and two gamma (γ) chains. There are two species of haemoglobin F, designated $^{G}\gamma$ and $^{A}\gamma$, with glycine and alanine, respectively, at position 136 of the γ chain. In addition, the $^{A}\gamma$chain shows polymorphism at position 75, which may be occupied by threonine

HAD1 8/5/05 2:11 PM Page 4

rather than the more common isoleucine [6], a polymorphism that was previously referred to as haemoglobin F-Sardinia. In the early embryo, haemoglobin is synthesized in the yolk sac and specific embryonic haemoglobins are produced—Gower 1, Gower 2 and Portland (or Portland 1). They contain globin chains that are synthesized in significant amounts only during embryonic life, specifically zeta (ζ) and epsilon (ε) chains (Table 1.1). Haemoglobins Gower 1 ($\zeta_2\varepsilon_2$) and Gower 2 ($\alpha_2\varepsilon_2$) were first described by Huehns and colleagues in 1961 [7], being named for Gower Street, in which University College Hospital is situated. Portland 1 ($\zeta_2\gamma_2$) was described in 1967 and was so named because it was first identified in the University of Oregon in Portland, Oregon [8]. By 5 weeks of gestation, ζ and ε chains are already being synthesized in primitive erythroblasts in the yolk sac. From the sixth week onwards, these same cells start to synthesize α, β and γ chains. Starting from about the 10th–12th week of gestation, there is haemoglobin synthesis in the liver and the spleen with the production of fetal and, later, adult haemoglobin. Production of the various embryonic, fetal and adult haemoglobins is synchronous in different sites. Later in intrauterine life the bone marrow takes over as the main site of haemoglobin synthesis and increasing amounts of haemoglobin A are produced. In adult life, bone marrow erythroblasts synthesize haemoglobin A and the minor haemoglobins.

The embryonic haemoglobins have a higher oxygen affinity than haemoglobin A, similar to that of haemoglobin F [9]. They differ from haemoglobins A and F in that they continue to bind oxygen strongly, even in acidotic conditions [9]. In the case of Gower 2, impaired binding to 2,3-diphosphoglycerate (2,3-DPG) is the basis of the increased oxygen affinity [10].

Haemoglobin can undergo post-translational modification (see also Chapter 6). Glycosylation occurs with the formation of haemoglobins $A_{1a–e}$, but principally haemoglobin A_{1c}. In normal individuals, haemoglobin A_{1c} may constitute up to 4–6% of total haemoglobin, but in diabetics it can be much higher. In individuals with a shortened red cell life span it is lower. Another minor fraction, formed on ageing, is haemoglobin A_{III}, in which glutathione is bound to the cysteine at β93. Unmodified haemoglobin can be distinguished by use of the designation haemoglobin A_0. In the fetus, about 20% of haemoglobin F shows acetylation of the γ chain, but this is not a major feature of other normal human globin chains [6]. Exposure to carbon monoxide, the product of incomplete combustion of hydrocarbons, leads to the formation of carboxyhaemoglobin. In normal individuals, carboxyhaemoglobin comprises 0.2–0.8% of total haemoglobin, but, in heavy smokers, it may be as much as 10–15%. Small amounts of sulphaemoglobin and methaemoglobin are also formed in normal subjects. Methaemoglobin is usually less than 1% of total haemoglobin. Post-synthetic modification of a haemoglobin molecule can also occur as a consequence of a mutation in a globin gene; either the abnormal amino acid or an adjacent normal amino acid can undergo post-translational conversion to another amino acid (see below). In addition, some abnormal haemoglobins, in which there is a mutation of N-terminal amino acids, are particularly prone to acetylation, which occurs cotranslationally [11].

The structure of haemoglobin is highly complex and can be viewed at four levels.

1 The primary structure is the sequence of the amino acids in the polypeptide which constitutes the globin chain.

2 The secondary structure is the arrangement of the polypeptide globin chains into α helices separated by non-helical turns; in the case of the β globin chain, there are eight α helices, designated A–H, whereas the α globin chain lacks the D helix residues; 70–80% of the amino acid residues of haemoglobin form part of the helices.

3 The tertiary structure is the arrangement of the coiled globin chain into a three-dimensional structure which has a surface haem-containing pocket between the E and F helices; binding of haem between two specific histidine residues in the E and F helices, respectively (Fig. 1.4), is essential for maintaining the secondary and tertiary structure of haemoglobin.

4 The quaternary structure is the relationship between the four globin chains, which is not fixed; the strong $\alpha_1\beta_1$ and $\alpha_2\beta_2$ bonds (dimeric bonds) hold the molecule together in a stable form, while the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ bonds (tetrameric bonds) both contribute to stability, albeit to a lesser extent than the dimeric bonds, and permit the chains to slide on each other and rotate; alteration in the quaternary structure of haemoglobin is responsible for the sigmoid oxygen

HAD1 8/5/05 2:11 PM Page 5

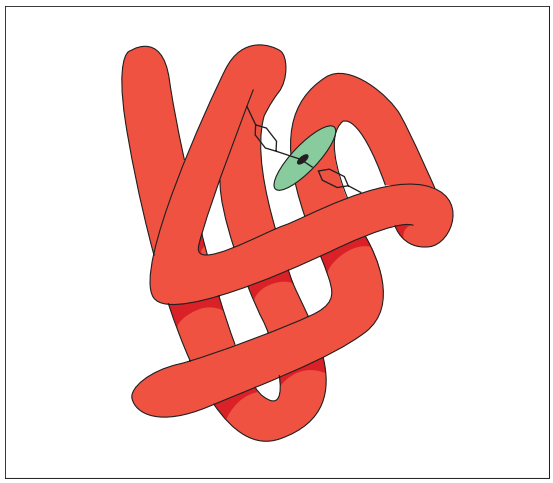

Fig. 1.4 Diagrammatic representation of a haemoglobin molecule with a haem group within the haem pocket, showing the relationship of haem to two histidine residues of the globin chain (designated proximal and distal histidines).

dissociation curve, the Bohr effect and the variation of oxygen affinity consequent on interaction with 2,3-DPG (see below). Contacts between like chains, $\alpha_1\alpha_2$ and $\beta_1\beta_2$, are also of physiological significance.

The interaction between the four globin chains is such that oxygenation of one haem group alters the shape of the molecule in such a way that oxygenation of other haem groups becomes more likely. This is known as cooperativity and is reflected in the shape of the oxygen dissociation curve (Fig. 1.5). The cooperativity between the globin chains is shown diagrammatically in Fig. 1.6. It is consequent on the fact that, in the deoxygenated state, the Fe^{2+} atom is out of the plane of the porphyrin ring of haem. Oxygenation of Fe^{2+} causes it to move into the plane of the porphyrin ring and, because of the link between haem and the histidine residues of globin, there is an alteration in the tertiary structure of that haemoglobin monomer; this, in turn, causes the oxygenated monomer to alter its position in relation to other

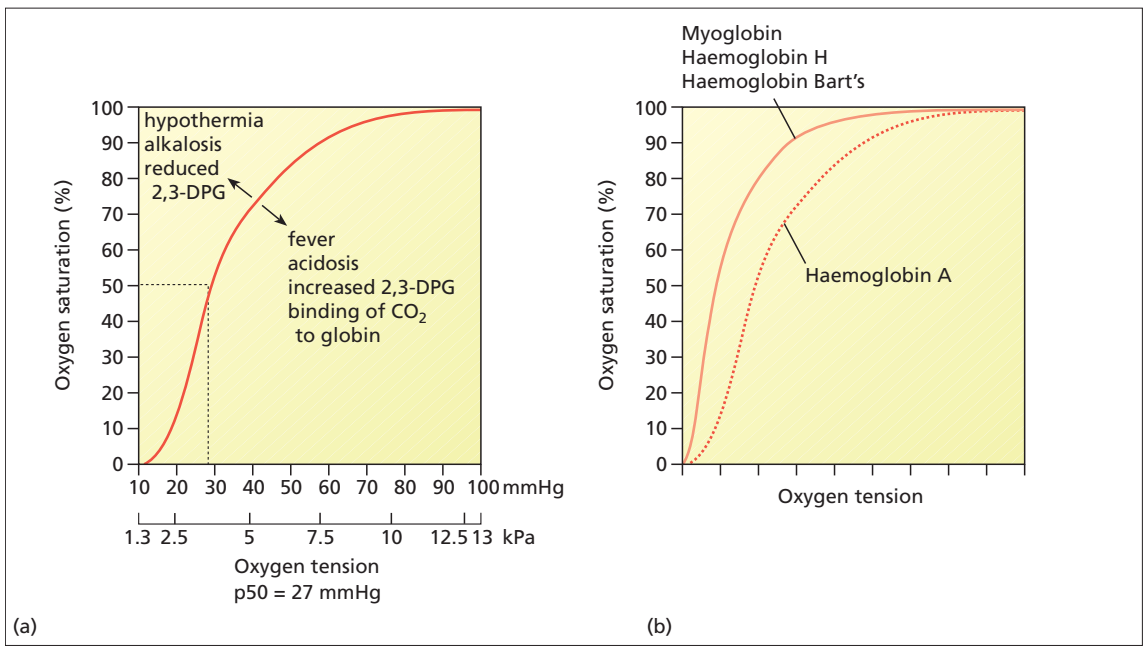

Fig. 1.5 (a) Normal oxygen dissociation curve indicating the effects of alteration of pH, body temperature and 2,3-diphosphoglycerate (2,3-DPG) concentration on the oxygen affinity of haemoglobin. (b) Comparison of the hyperbolic oxygen dissociation curve characteristic of myoglobin and of abnormal haemoglobins that do not exhibit cooperativity, with the sigmoid dissociation curve characteristic of haemoglobin A; haemoglobins A_2 and F have dissociation curves similar to that of haemoglobin A but further to the right.

HAD1 8/5/05 2:11 PM Page 6

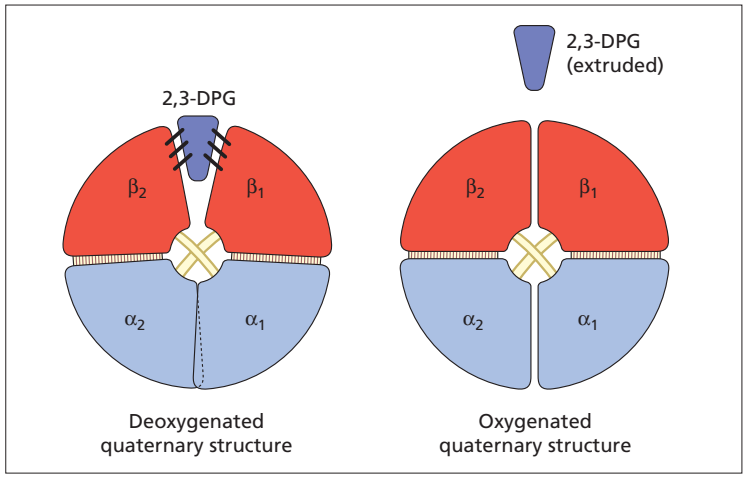

Fig. 1.6 Diagrammatic representation of the effect of oxygenation and deoxygenation on the quaternary structure of haemoglobin. The haemoglobin dimers ($\alpha_1\beta_1$ and $\alpha_2\beta_2$) are stable, with the dimeric bonds between the α and β chains having 34 contacts in both the deoxygenated and oxygenated forms. There are less strong $\alpha_2\beta_1$ and $\alpha_1\beta_2$ tetrameric bonds, with 17 contacts between the α and β chains in the deoxy form and a different 17 contacts in the oxy form. There are also $\alpha_1\alpha_2$ bonds with four inter-chain contacts in the deoxy form only. 2,3-diphosphoglycerate (2,3-DPG) binds to the β chains (three contacts with each chain) only in the deoxy form of the molecule. Oxygenation is associated with breaking and reforming of tetrameric ($\alpha_2\beta_1$ and $\alpha_1\beta_2$) contacts, breaking of $\alpha_1\alpha_2$ contacts, expulsion of 2,3-DPG and the assumption of a more compact form of the molecule. In the deoxygenated form, the α chains are closer together and there is a cleft between the β chains, whereas, in the oxygenated form, the α chains are further apart and the β cleft has disappeared.

haemoglobin monomers, i.e. the quaternary structure of the haemoglobin molecule is altered. The oxygenated haemoglobin molecule is smaller than the non-oxygenated molecule. Cooperativity between the globin chains is also the basis of the alkaline Bohr effect (often referred to simply as the Bohr effect), i.e. the reduction in oxygen affinity that occurs when the pH falls from physiological levels of 7.35–7.45 towards 6.0. Increasing metabolism in tissues lowers the pH as there is increased production of CO_2 and carbonic acid and, in addition, in anaerobic conditions, the generation of lactic acid. The Bohr effect therefore leads to enhanced delivery of oxygen to tissues, such as exercising muscle. Similarly, the quaternary structure of haemoglobin makes possible the interaction of haemoglobin with 2,3-DPG, which enhances oxygen delivery. Synthesis of 2,3-DPG is increased by hypoxia. Marked anaemia can cause respiratory alkalosis, which enhances 2,3-DPG synthesis, thus compensating to some extent for the anaemia. There is also increased 2,3-DPG synthesis in renal failure, again partly compensating for the anaemia.

Oxygen affinity is reduced not only by acidosis and increased levels of 2,3-DPG, but also by fever. All of these effects are likely to be of physiological significance. Fever increases the metabolic rate, so that decreased oxygen affinity, favouring downloading of oxygen, is beneficial in this circumstance. The lower pH in tissues favours the delivery of oxygen to sites of active metabolism, whereas the efflux of CO_2 in the lungs raises the pH and favours the uptake of oxygen by haemoglobin. It should be noted that the acute effect of acidosis and the chronic effect of respiratory alkalosis both contribute to improved oxygen delivery to tissues.

Genetics of haemoglobin synthesis

Haem synthesis takes place in erythroid precursors, from the proerythroblast stage to the reticulocyte stage. Eight enzymes, under separate genetic control, are known to be necessary for haem synthesis [12]. Different stages of haem synthesis take place either in mitochondria or within the cytosol (Fig. 1.7). The first enzymatic reaction and the last three occur

HAD1 8/5/05 2:11 PM Page 7

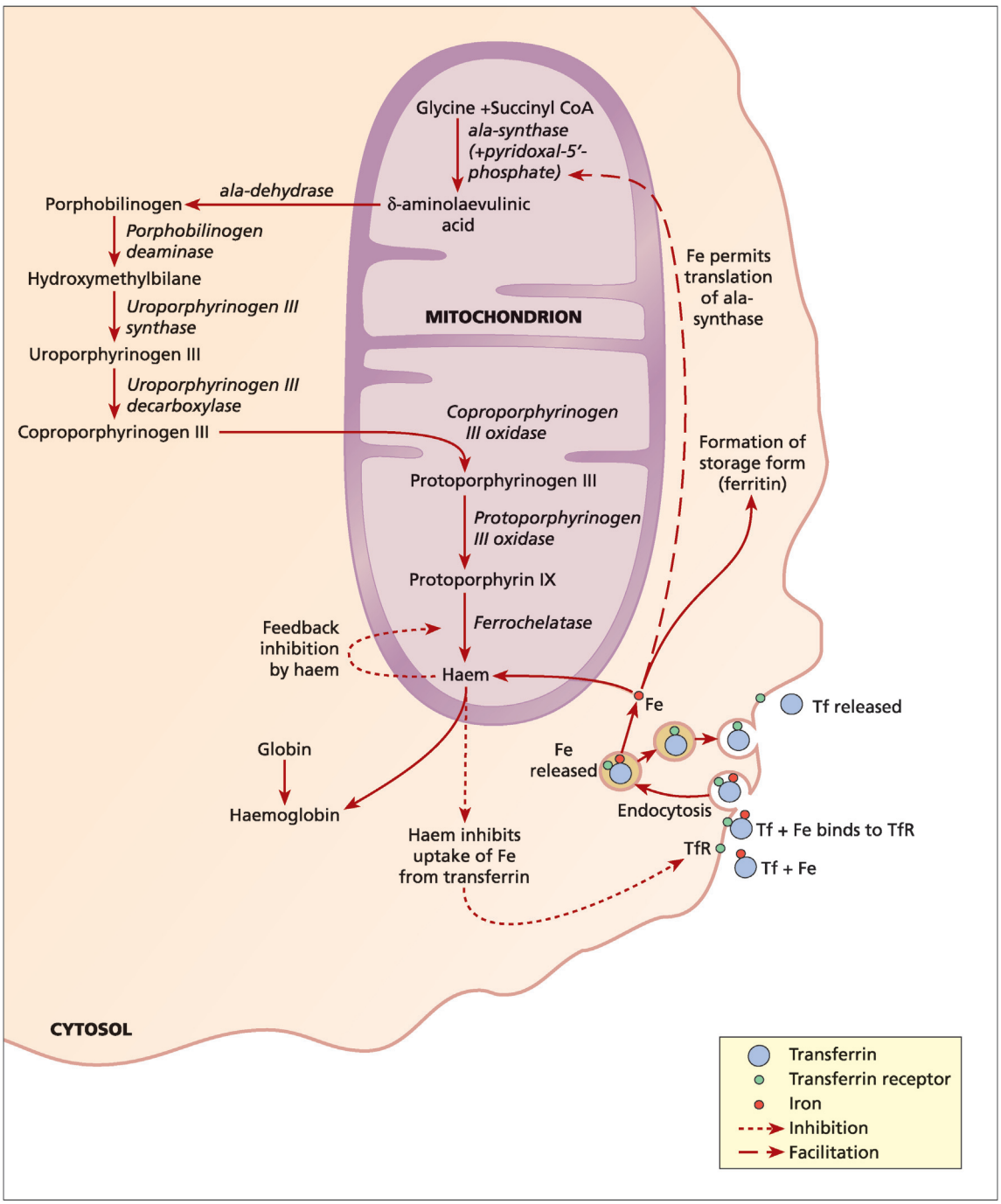

Fig. 1.7 Diagrammatic representation of haem synthesis.

HAD1 8/5/05 2:11 PM Page 8

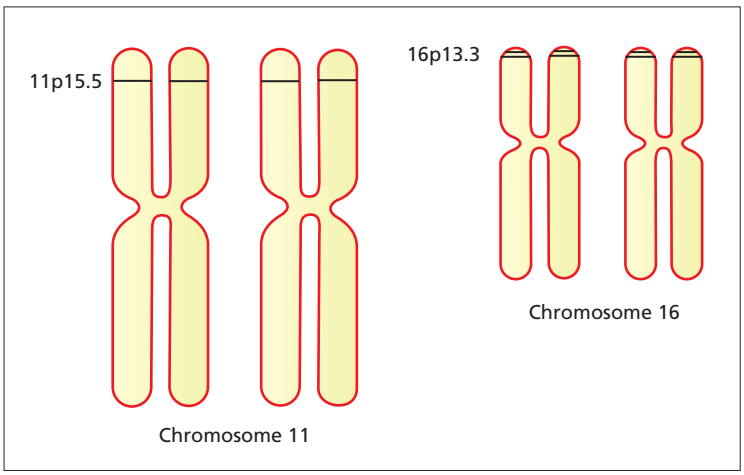

Fig. 1.8 Diagram of chromosomes 11 and 16 showing the positions of the β and α globin gene clusters.

in the mitochondrion, whereas the four intermediate enzymatic reactions occur in the cytosol. The first rate-limiting step in haem synthesis is the formation of δ-aminolaevulinic acid (ala) by condensation of glycine and succinyl CoA. This reaction is under the control of ala-synthase with pyridoxal-5′-phosphate as cofactor. The rate of formation of ala is controlled by iron availability; iron deficiency causes iron regulatory proteins to bind to iron-responsive elements in the messenger RNA (mRNA) for ala-synthase with consequent repression of translation. Synthesis of ala is followed by its entry into the cytosol where two molecules combine, under the influence of ala-dehydrase, to form porphobilinogen. Four molecules of porphobilinogen in turn combine to form uroporphyrinogen III, which is then modified in two further steps to form coproporphyrinogen III. Coproporphyrinogen III enters the mitochondrion where it is converted to protoporphyrin IX. The final stage is the combination of ferrous (Fe^{2+}) iron with protoporphyrin IX to form haem, under the influence of ferrochelatase. Haem is also referred to as ferroprotoporphyrin.

The uptake of iron by erythroid cells is from transferrin (Fig. 1.7). A molecule of transferrin with its attached iron first binds to a membrane transferrin receptor. The whole complex is internalized, a process known as endocytosis. Iron is released from its carrier within the endocytotic vesicle and, following reduction to the ferrous form, is transferred to the mitochondrion for haem synthesis or is stored as ferritin within the cytoplasm. The transferrin molecule then detaches from the transferrin receptor and is released from the cell surface. There is negative feedback control of haem synthesis by haem, which inhibits both ferrochelatase and the acquisition of iron from transferrin. Reduced cellular uptake of iron in turn inhibits the production of ala. Uptake of iron by erythroid cells is enhanced by iron deficiency and by increased levels of erythropoietin. Both lead to the combination of iron regulatory proteins with iron-responsive elements in the mRNA for the transferrin receptor protein. The mRNA is then protected from degradation, leading to increased expression of transferrin receptors on erythroid cell membranes and increased iron uptake.

The synthesis of α and β globin chains takes place in erythroid precursors, from the proerythroblast to the reticulocyte stage; δ chain synthesis ceases before the reticulocyte stage [13]. Synthesis is in the cytoplasm, on ribosomes. The genes for globin chain synthesis are located in two clusters, on chromosomes 11 and 16 (Figs 1.8 and 1.9). The α gene cluster is close to the telomere of chromosome 16, at 16p13.3. The distance from the telomere shows polymorphic variation, from 170 to 430 kilobases (kb), a kilobase being 1000 nucleotide bases. The β gene is at 11p15.5. In addition to the functional globin genes, these clusters contain 'pseudogenes', which are non-functional homologues of globin genes; they are transcribed but not translated. The α cluster of chromosome 16 extends over 28 kb and contains, in the following order,

HAD1 8/6/05 9:28 AM Page 9

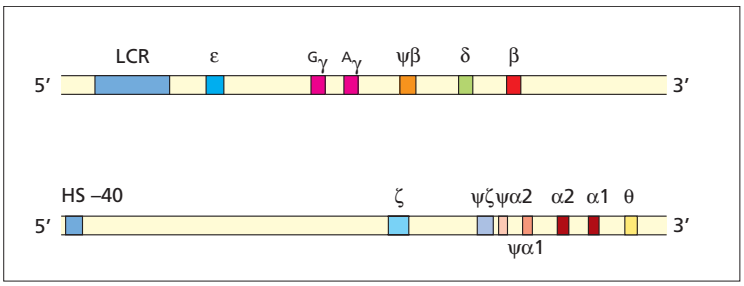

Fig. 1.9 Diagrammatic representation of the α and β globin gene clusters.

Table 1.2 Sequences showing CACCC, CCAAT and TATA homology in the promoters of globin genes; identical sequences in different genes are shown in bold red.

Gene	CACCC homology box	CCAAT homology box	TATA homology box
ζ		**CCAAT**	T**ATA**AAC
α1 and α2		**CCAAT**	C**ATA**AAC
ε		**CCAAT**	A**ATA**AAG
$^{G}\gamma$ and $^{A}\gamma$	**CACCC**	**CCAAT/CCAAT**	A**ATA**AAA
β	**CACCC**	**CCAAT**	C**ATA**AAA
δ		**CCAA**C	C**ATA**AAA

a ζ gene (also referred to as ζ2), a pseudo ζ gene (ψζ or ψζ1), two pseudo α genes, (ψα2 and ψα1) and two α genes, designated α2 and α1. The β cluster on chromosome 11 contains, in the following order, an ε gene, two γ genes, designated $^{G}\gamma$ and $^{A}\gamma$, respectively, a pseudo β gene (ψβ), a δ gene and a β gene. There is wide variability of the α and β globin gene clusters between individuals and groups, with duplications and triplications of ζ, ψζ and α being quite common. The overall structure of the two clusters is remarkably conserved amongst vertebrates and this has led to the hypothesis that all the globin genes, as well as the gene for the unlinked but related protein, myoglobin, arose from a common ancestor by the processes of duplication, unequal crossing over and sequence divergence. Many primitive invertebrates have only a single globin gene, whereas fish and amphibians have α and β genes on the same chromosome. Birds have α and β genes on different chromosomes. All the human globin genes have three coding sequences (exons) and two intervening non-coding sequences (intervening sequences or introns) and are flanked by 5′ and 3′ non-coding sequences (referred to as untranslated regions, UTRs) (Fig. 1.10). The two α genes differ in structure in intron 2 and the 3′ UTR, but the coding sequences are identical. As for all genes, coding is by means of triplets of nucleotides, known as codons, which code for a specific amino acid. 5′ to each gene is the promoter, a sequence that binds RNA polymerase and transcription factors and is necessary for the initiation of transcription. Globin gene promoters share several conserved DNA sequences that bind crucial transcription factors [14,15]. These are summarized in Table 1.2.

The process by which globin chains are synthesized is shown diagrammatically in Fig. 1.10. Transcription is the process by which RNA is synthesized from a DNA template by the action of RNA polymerase. The entire globin gene, including the introns and the 5′ and 3′ UTRs, is transcribed. Transcription is controlled by interaction between the genes and transcription factors that bind to both promoters and upstream regulatory elements, referred to as the β-locus control region (β-LCR) for the β cluster and HS –40 for the α cluster. The β-LCR includes four

HAD1 8/5/05 2:11 PM Page 10

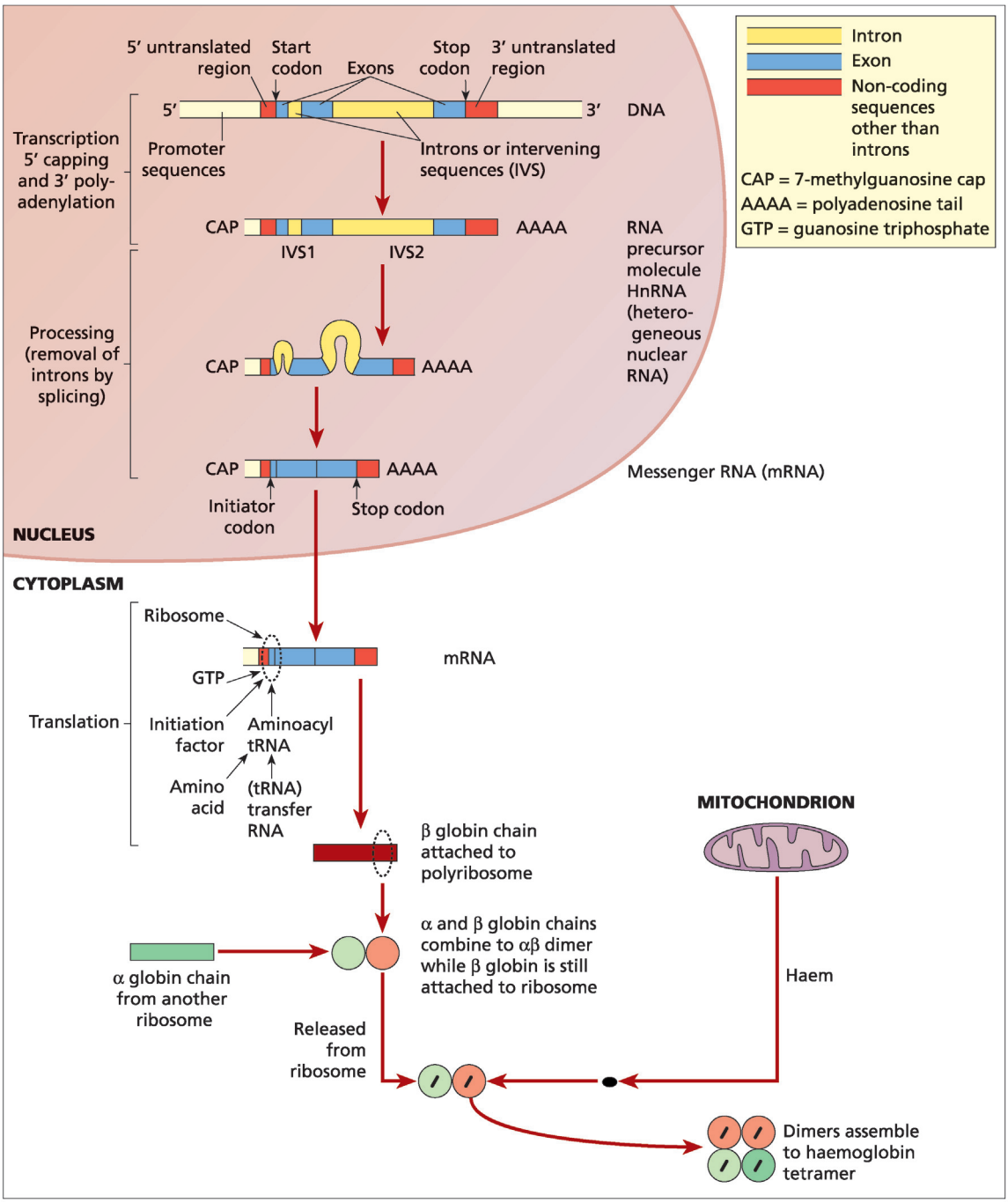

Fig. 1.10 Diagrammatic representation of RNA synthesis and processing and β globin chain synthesis.

HAD1 8/5/05 2:11 PM Page 11

erythroid-specific DNase sites, designated HS1, HS2, HS3 and HS4, of which HS3 is probably the most important in opening the chromatin structure to permit access of transcription factors and HS2 is probably the most important in enhancing globin chain synthesis [16]. There are also enhancers within introns of genes and downstream of the β and $^{A}\gamma$ genes. *Trans*-acting factors, encoded by genes on chromosomes other than 11 and 16, are vital for the expression of globin genes. Relatively erythroid-specific *trans*-activating factors, including GATA1, NFE2, EKLF, SSP, Nrf-1, Nrf-2 and LCR-F1, contribute to the regulation of gene expression by interacting either with the LCRs or with the globin gene promoters to increase gene expression [17,18]. EKLF (erythroid Kruppel-like factor) is an enhancer of β chain synthesis and SSP (stage selector protein) is an enhancer of δ and γ chain synthesis [17]. In addition to transcription factors that are relatively specific to erythroid cells, globin gene expression is also influenced by general transcription factors, including AP-1, Sp1, YY1, USF and TAL-1/SCL [16–18]. Nascent RNA molecules resulting from transcription are large and unstable and are modified in the nucleus. Initially, the 5′ end acquires a 7-methylguanosine cap (CAP), which is probably added during transcription and has a role during translation; during this 'capping' process, methylation of adjacent ribose residues also occurs. Following this, the majority of transcripts acquire a 3′ polyadenosine tail with the addition of 75 to several hundred adenylate residues. There is an AAUAA sequence near the 3′ end (within the 3′ UTR) that serves as a signal for 3′ cleaving of the transcript and polyadenylation. Polyadenylation may have a role in transfer of the mRNA from the nucleus to the cytoplasm. The polyadenylate tail is also important for mRNA stability and enhances translation. Finally, the introns are excised to give a functional mRNA molecule which, in most cases, contains a single continuous open reading frame (ORF), encoding the sequence of the relevant protein, flanked by 5′ and 3′ UTRs.

Molecules of mRNA move from the nucleus to the cytoplasm where they bind to ribosomes and serve as templates for the assembly of the polypeptide sequences of the globin chain. Each nucleotide triplet serves as a template for a specific amino acid that is covalently bound to, and transported to the ribosome by, transfer RNA (tRNA). tRNAs are specific for both a nucleotide triplet and an amino acid. Amino acids are thus assembled in the correct sequence, forming a polypeptide. This process is known as translation. An initiation codon, AUG, is essential for the initiation of translation; it is the first codon after the 5′ UTR and encodes methionine. Initiation requires the amino acid methionine, tRNA specific for methionine, guanosine triphosphate (GTP) and an initiation factor. When the nascent molecule reaches 20–30 amino acid residues, the methionine is removed through the action of methionine aminopeptidase. When the chain reaches 40–50 residues, cotranslational acetylation of the N-terminal residue can occur through the action of several acetyl transferases [19]. Whether this occurs to any great extent depends on the nature of the N-terminal residue. Thus the glycine of the γ chain is 10–15% acetylated, whereas the valine of normal α, β and δ chains is resistant to acetylation. There are 64 possible nucleotide triplets or codons, 61 of which encode amino acids (20 in all) and three of which do not; the latter serve as stop or termination codons, leading to termination of globin chain synthesis. Transcription thus continues until a termination codon, UAA, UAG or UGA, is encountered. The termination codon is followed by the 3′ UTR.

The rate-limiting step of globin chain translation is the commencement of elongation, i.e. the next step after initiation. Transcription from the two α genes is equal up to the eighth week of gestation, but thereafter the α2 gene becomes dominant and, in adult life, the ratio of α2 to α1 mRNA is 2.6–2.8 : 1 [20]. The translational efficiency differs somewhat so that the α2 gene directs the synthesis of about twice as much α chain as the α1 gene. There is more α than β mRNA, probably about 2.5 times as much, but β chain synthesis is more translationally efficient than α chain synthesis and α chains are therefore produced only slightly in excess of β chains [20]. The control of globin chain synthesis is probably mainly at the level of transcription, with translational control being less important. Translation is dependent on the presence of haem. In iron deficiency, the reduced availability of haem leads to inactivation of the initiation factor and thus reduced synthesis of globin chains. The α and β globin chains are synthesized on different polyribosomes. The combination of a free α

HAD1 8/5/05 2:11 PM Page 12

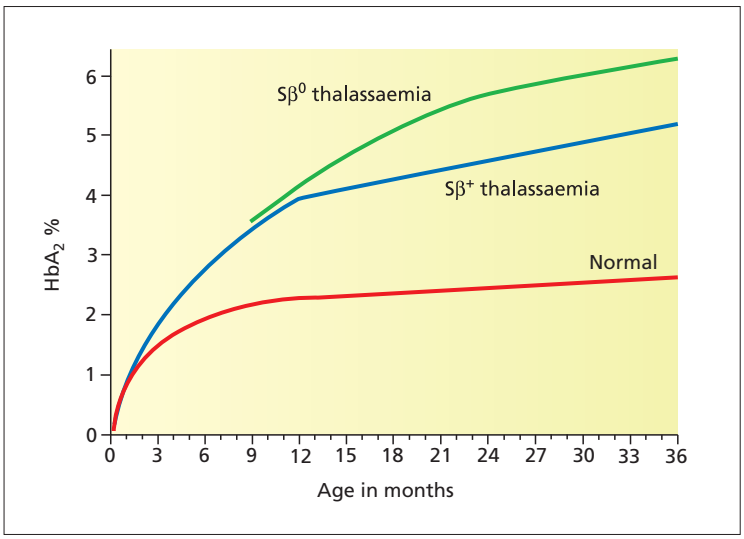

Fig. 1.11 Diagram showing the rate of rise of haemoglobin A_2 in haematologically normal Jamaican babies and in babies with sickle cell/β thalassaemia. (Modified from reference [21].)

chain with a β chain that is still attached to the polyribosome, to form an αβ dimer, may contribute to the release of the β chain from the ribosome. Incorporation of haem probably occurs after release from the polyribosome.

Globin mRNA is unusually stable so that translation can continue for up to 3 days after cessation of transcription. Both the α and β globin genes have structural determinants in their 3′ UTRs that are important for mRNA stability [17].

Normal haemoglobins

The normal haemoglobins beyond the neonatal period are haemoglobin A and two minor haemoglobins, haemoglobin A_2 and haemoglobin F.

Haemoglobin A_2

In adults, haemoglobin A_2 comprises about 2–3.5% of total haemoglobin. The percentage is much lower at birth, about 0.2–0.3%, with a rise to adult levels during the first 2 years of life. The steepest rise occurs in the first year, but there is a continuing slow rise up to 3 years of age [21] (Fig. 1.11). In the normal adult population, the percentage of haemoglobin A_2 shows a Gaussian distribution. It has functional properties that are very similar to those of haemoglobin A [13] (similar cooperativity and interaction with 2,3-DPG), although, in comparison with haemoglobin A, it inhibits polymerization of haemoglobin S [22] and has a higher oxygen affinity [10]. It has a pancellular distribution.

The reduced rate of synthesis of haemoglobin A_2, in comparison with haemoglobin A, reflects the much slower rate of synthesis of the δ chain in comparison with the β chain. This, in turn, appears to be consequent on a reduced rate of transcription of δ mRNA caused by a difference in the promoter region of these two genes; the δ gene has a CCAAC box rather than the CCAAT box of the β gene [13] and, in addition, lacks the CACCC sequence that is present in the β promoter (Table 1.2). The proportion of haemoglobin A_2 is reduced by absolute or functional iron deficiency (see Table 6.3) and by α, δ and δβ thalassaemia trait (see Fig. 3.11). In γδβ thalassaemia, the rate of synthesis, but not the proportion, of haemoglobin A_2 is reduced, as the synthesis of γ and β chains is reduced, as well as δ chain synthesis. The proportion of haemoglobin A_2 is increased in the great majority of patients with β thalassaemia trait and in some patients with an unstable haemoglobin.

There are δ chain variants and δ thalassaemias. About 1% of individuals of African ancestry have the variant haemoglobin designated haemoglobin A_2' (A_2 prime) or haemoglobin B_2 ($\delta^{16Gly\rightarrow Arg}$). It is readily detected by high performance liquid chromatography (Fig. 1.12) and isoelectric focusing. The δ thalas-

HAD1 8/5/05 2:11 PM Page 13

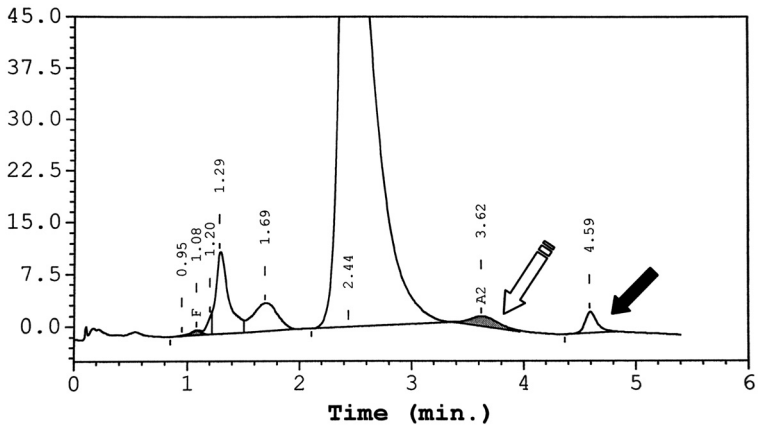

Fig. 1.12 High performance liquid chromatography (HPLC) chromatogram showing a split haemoglobin A_2 resulting from heterozygosity for haemoglobin A_2'; the white arrow shows haemoglobin A_2 and the black arrow haemoglobin A_2'.

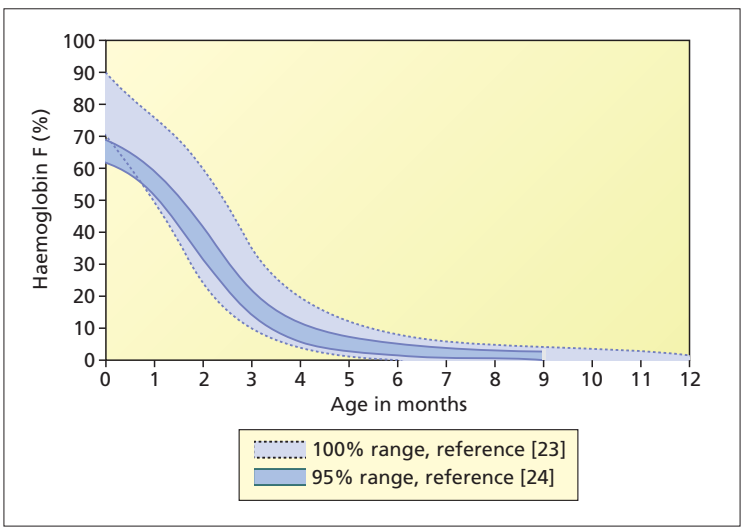

Fig. 1.13 Rate of fall of the percentage of haemoglobin F postnatally in normal and premature babies; the pale blue represents premature babies while the deep blue represents normal babies. (Derived from references [23,24].)

saemias are also common in some ethnic groups, e.g. present in 1% of Sardinians [10]. The δ thalassaemias and δ chain variants are of no functional significance, although some variants are unstable or have increased oxygen affinity. However, their presence complicates the diagnosis of β thalassaemia trait (see p. 97).

Haemoglobin F

Haemoglobin F is the major haemoglobin during intrauterine life. Its oxygen affinity is higher than that of haemoglobin A and this facilitates oxygen transfer from the mother to the fetus. However, it should be noted that fetal development appears to be normal in the offspring of mothers with very high levels of haemoglobin F. Its oxygen dissociation curve is sigmoid. The increased oxygen affinity, in comparison with haemoglobin A, is attributable to its weak affinity for 2,3-DPG [6]. In comparison with haemoglobin A, haemoglobin F is less efficient at transporting CO_2. A significant proportion of haemoglobin F is acetylated.

During the first year of life, the percentage of haemoglobin F falls progressively to values close to adult levels (Fig. 1.13) [23–25]. A slower fall to final adult levels may continue for several years, even up

HAD1 8/5/05 2:11 PM Page 14

to puberty and beyond. The percentage of fetal haemoglobin present at birth is quite variable, usually being between 60% and 95%. During intrauterine life and at birth, haemoglobin F shows a $^{G}\gamma$ to $^{A}\gamma$ ratio of approximately 2:1 to 3:1. Within the first few months of birth, this changes to the adult ratio of approximately 2:3. In premature infants, there is initially a plateau phase in haemoglobin F concentration lasting 20–60 days, followed by a linear decrease similar to that in term babies [24]. At any given period after birth, the spread of values is greater than that in term babies. Initially, there are more high values but, after the first month of life, values both higher and lower than those of term infants are observed [24].

In normal adults, haemoglobin F is heterogeneously distributed, being found in a subset of erythrocytes designated F cells. The proportion of F cells is highly variable, in one study ranging from 0.6% to 22% [26]. The percentage of haemoglobin F is determined by age, sex (slightly higher in women) and a number of inherited characteristics both linked and unlinked to the β globin gene cluster. DNA sequences controlling the proportion of F cells and the percentage of haemoglobin F include [18,23–29]:

- a polymorphism at position –158 of the Gγ gene (C→T being associated with a higher haemoglobin F);
- variation of the number of repeats of a specific motif at –530 in the HS2 component of the β-LCR, namely $(AT)_xN_{12}GT(AT)_y$;
- a *trans*-acting locus at 6q22.3–23.2;
- a *trans*-acting locus at Xp22.2–22.3;
- a *trans*-acting locus on an autosome other than 6.

The percentage of haemoglobin F is also affected by any increase in the number of γ genes.

The mechanism by which the polymorphisms in the LCR at –530 base pairs (bp) to the Gγ gene influence γ chain synthesis appears to be that, in comparison with $(AT)_7T_7$, the $(AT)_9T_5$ sequence shows increased binding of BP-1, a negative *trans*-acting factor [30].

The distribution of the percentage of haemoglobin F in the population is skewed. In 85–90% of individuals, haemoglobin F is less than 0.6–0.7% and F cells are less than 4.5% [27,29]. The other 10–15% of the population have values above these levels. The upper limit of normal is rather arbitrarily taken as 1%. It would probably be more accurate to take 0.6% or 0.7% as the upper limit of normal, excluding the 11% of males and 21% of females who have a slight elevation of the percentage of F cells and the haemoglobin F percentage as an X-linked dominant characteristic [27]. However, as the measurement of a low percentage of haemoglobin F is very imprecise, 1% is a practical upper limit.

Haemoglobin F is more markedly increased in patients with various inherited abnormalities of β globin chain synthesis (see Table 3.12) and, less often, in various acquired conditions (see Table 6.2).

Variant haemoglobins and abnormalities of globin gene synthesis

Nuclear DNA, including the DNA of globin genes, is subject to spontaneous mutation. This may be a point mutation (alteration of a single nucleotide) or a more extensive mutation, in which there is deletion, insertion or other alteration of more than one nucleotide. The types of mutation that may occur in globin genes are summarized in Table 1.3. In addition, expression of globin genes can be affected by DNA sequences outside the globin genes themselves, either enhancers acting in *cis* or genes on other chromosomes encoding *trans*-acting transcription factors (Tables 1.3 and 1.4).

Point mutations in globin genes sometimes have no effect on the amino acid sequence. This occurs because, as mentioned above, there is redundancy in the genetic code, with a number of nucleotide triplets coding for the same amino acid. When a 'same-sense' mutation occurs, the new codon resulting from the mutation codes for the same amino acid as the original codon and there is thus no effect on the final gene product. Similarly, mutation of a termination codon may be to a different termination codon. Many spontaneous mutations in globin genes are same-sense mutations. Point mutations may also result in a 'missense' mutation when the new codon codes for a different amino acid, leading to the production of a variant haemoglobin. The site of a mutation is critical, determining whether there is an effect on stability, oxygen affinity, solubility or other critical characteristics of the haemoglobin molecule. Because of the redundancy in the genetic code, different point mutations may give rise to the same variant

HAD1 8/5/05 2:11 PM Page 15

Table 1.3 Types of mutation that can occur in globin genes and adjoining sequences.

Type of mutation	Possible consequence	Example
Point mutations		
Within coding sequence, i.e. within an exon	**Same-sense or neutral mutation**, i.e. mutant codon codes for same amino acid as normal codon, so there are no consequences	Many mutations are of this type; more than one-third of theoretically possible point mutations would result in no alteration in the amino acid encoded
	Mis-sense mutation, i.e. mutant codon codes for a different amino acid from the normal codon; includes mis-sense mutations in which an abnormal amino acid interferes with the normal cleavage of the N-terminal methionine	Haemoglobin S, haemoglobin C, haemoglobin E Haemoglobin Marseille and haemoglobin South Florida (altered amino acid near N-terminus plus persisting methionine residue at the N-terminus of the β chain)
	Nonsense mutation, i.e. the mutant codon does not code for an amino acid and thus functions as a stop or termination codon, producing a shortened globin chain	Haemoglobin McKees Rocks (two amino acids shorter than normal); α2 CD116 GAG→TAG creating premature stop codon and causing α thalassaemia
	New-sense mutation, i.e. conversion of a stop codon to a coding sequence, producing an elongated globin chain	Haemoglobin Constant Spring, haemoglobin Icaria, haemoglobin Seal Rock, haemoglobin Koya Dora, haemoglobin Paksé
	Gene conversion*	Conversion of $^{G}\gamma$ gene to $^{A}\gamma$ gene, giving $^{A}\gamma^{A}\gamma$ genotype Conversion of $^{A}\gamma$ gene to $^{G}\gamma$ gene, giving $^{G}\gamma^{G}\gamma$ genotype Conversion of ψζ1 to a gene that resembles ζ2 but is still non-functional (ζ1) Conversion between the α2 and α1 genes so that the same mutation is present in both, e.g. $\alpha2^{Lys\rightarrow Glu}\alpha1^{Lys\rightarrow Glu}$, giving unusually high levels of haemoglobin I
	Gene conversion plus further point mutation*	Haemoglobin F-Port Royal, resulting from a further point mutation in a $^{G}\gamma^{G}\gamma$ gene complex
Within non-coding sequence, i.e. in an intron	Production of a new splice site leading to a structurally abnormal mRNA	Some β thalassaemias
Mutation 5′ or 3′ to the gene (i.e. outside the gene)	Mutation of an enhancer	Some β thalassaemias
	Reduced rate of synthesis of mRNA due to interference with 3′-end formation of mRNA	Some β thalassaemias

Continued on p. 16.

HAD1 8/5/05 2:11 PM Page 16

Table 1.3 *Continued.*

Type of mutation	Possible consequence	Example
Deletion or duplication of one or more genes		
Deletion of one or more genes	Total loss of expression of relevant gene; occasionally, also loss of function of an adjacent structurally normal gene	Most α thalassaemias, some β thalassaemias, δβ thalassaemias and γδβ thalassaemias; deletion of $^{G}\gamma$ gene ($-^{A}\gamma$), homozygosity for which causes anaemia and a reduced haemoglobin F percentage in the neonate; deletion of ψζ1
Deletion of genes with downstream enhancer being juxtaposed to remaining gene	Loss of β and δ gene function, but enhanced function of remaining $^{G}\gamma$ ($\pm^{A}\gamma$) gene	Deletional hereditary persistence of fetal haemoglobin
Duplication of α gene	Triple or quadruple α gene	ααα†/αα, ααα/ααα, αααα‡/αα or αααα/αααα
Triplication of entire α globin gene cluster	Six α genes on a single chromosome	αα:αα:αα/αα
Duplication of $^{G}\gamma$ gene	Double, triple or quadruple $^{G}\gamma$ gene so that there are three, four or five γ genes on a chromosome	$^{G}\gamma^{G}\gamma^{A}\gamma$, $^{G}\gamma^{G}\gamma^{G}\gamma^{A}\gamma$ (homozygotes have been described with a total of eight γ genes) or $^{G}\gamma^{G}\gamma^{G}\gamma^{G}\gamma^{A}\gamma$
Duplication of the ζ or ψζ gene	Double, triple or quadruple ζ/ψζ gene	ζ2ψζ1ψζ1/ζ2ψζ1 or ζ2ψζ1ψζ1/ζ2ψζ1ψζ1 or four ζ-like genes per chromosome
Abnormal cross-over during meiosis leading to gene fusion		
α2α1 fusion	Effective loss of one α gene but structurally normal α chain is encoded	$-\alpha^{3.7}$ thalassaemia
δβ fusion — simple cross-over	Reduced rate of synthesis of structurally abnormal globin chain	Haemoglobin Lepore, e.g. haemoglobin Lepore-Washington/Boston, haemoglobin Lepore-Baltimore and haemoglobin Lepore-Hollandia, or $\delta^{0}\beta^{+}$ thalassaemia [31]
δβδ fusion — double cross-over with δ sequences on either side of β sequences	Reduced rate of synthesis of structurally abnormal globin chain	Haemoglobin Parchman
βδ fusion (with preservation of intact δ and β genes on either side of fusion gene, with or without additional mutation)		Anti-Lepore haemoglobins, e.g. haemoglobin Miyada, haemoglobin P-Nilotic, haemoglobin P-Congo, haemoglobin Lincoln Park
$^{A}\gamma\beta$ fusion	Synthesis of variant haemoglobin plus increased synthesis of haemoglobin F	Haemoglobin Kenya
$\beta^{A}\gamma$ fusion (with preservation of intact $^{G}\gamma$ and $^{A}\gamma$ genes and duplication of the δ gene)		Haemoglobin anti-Kenya
$^{G}\gamma^{A}\gamma$ fusion (designated $-^{G}\gamma^{A}\gamma-$)	Reduced rate of synthesis of haemoglobin F	γ thalassaemia

HAD1 8/5/05 2:11 PM Page 17

Table 1.3 *Continued.*

Type of mutation	Possible consequence	Example
Deletion of DNA sequences but without a frame shift in coding sequence		
Deletion of part of a coding sequence, either three nucleotides or a multiple of three	One to five amino acids missing but sequence otherwise normal	Haemoglobin Gun Hill (an unstable haemoglobin with five amino acids missing)
Deletion plus inversion		
Two deletions with inversion of intervening sequence	Deletion involving $^{A}\gamma$ and δ plus β genes, respectively, but with preservation of an intervening region which is inverted	Indian type of deletional $^{A}\gamma\delta\beta^{0}$ thalassaemia
Deletion plus insertion		
Deletion with insertion of extraneous DNA between breakpoints	Same functional effect as deletion	One type of α^{0} thalassaemia, $--^{MED}$
Insertion within a coding sequence but without a frame shift		
Insertion of nucleotides, either three or multiples of three, e.g. by tandem duplication	Up to five extra amino acids	Haemoglobin Koriyama (an unstable haemoglobin with insertion of five codons in β gene, anti-Gun Hill); haemoglobin Grady (insertion of three codons in α gene)
Frame shift mutations		
Alteration of the reading frame resulting from deletion, insertion, deletion plus insertion or deletion plus duplication	Abnormal amino acid sequence with an elongated globin chain (when a stop codon is out of phase and translation continues until another 'in-frame' stop codon is met); abnormal amino acid sequence with a truncated globin chain (when a premature stop codon is created)	Haemoglobin Wayne (α chain), haemoglobin Tak (β chain), haemoglobin Cranston (β chain), some β thalassaemias, including some dominant β thalassaemias, some α thalassaemias
Chromosomal translocation		
Unbalanced translocation	Extra α genes on a chromosome other than chromosome 16	Same significance as homozygous triplication of an α gene as there are a total of six α genes
	Loss of an α gene	α thalassaemia
Deletion of a locus control region		
Locus control region deleted, with or without deletion of relevant genes	Deletion of the locus control region of the β gene	$(\varepsilon)\gamma\delta\beta^{0}$ thalassaemia
	Deletion of the α gene enhancer (HS –40) 40 kb upstream of the ζ2 gen	α^{0} thalassaemia

*Gene conversion is non-reciprocal genetic exchange between allelic or non-allelic homologous sequences so that one gene comes to resemble another more closely or becomes identical to it; it is responsible for maintaining the similarity between pairs of identical or similar genes.
† Either $\alpha\alpha\alpha^{anti3.7}$ or $\alpha\alpha\alpha^{anti4.2}$.
‡ Either $\alpha\alpha\alpha\alpha^{anti3.7}$ or $\alpha\alpha\alpha\alpha^{anti4.2}$.

HAD1 8/5/05 2:11 PM Page 18

Table 1.4 Mutations occurring outside the globin gene clusters leading to abnormal globin gene synthesis.

Mutation	Consequence
Mutation in XH_2 gene at Xq13.3, which encodes a *trans*-acting factor regulating α gene expression	Haemoglobin H disease plus dysmorphism and severe learning difficulties
Mutation in a putative gene at Xp22.2–22.3	Hereditary persistence of fetal haemoglobin
Mutation in a putative gene at 6q22.3–23.1	Hereditary persistence of fetal haemoglobin
Mutation in the *XPD* gene at 19q13.2–13.3, which encodes one component of the general transcription factor, TFIIH [32]	Recessive trichothiodystrophy and β thalassaemia
Mutation in the *GATA1* gene at Xp11–12 [33]	X-linked thrombocytopenia and β thalassaemia

haemoglobin. For example, the α chain variant, G-Philadelphia, has arisen twice, from an AAC to AAG change in an α2α1 fusion gene and from an AAC to AAA change in an α2 gene [34]. There are more than 900 known variant haemoglobins resulting from point mutations. Some point mutations are 'nonsense' mutations in which the new codon is one of the three that do not code for an amino acid. A nonsense mutation thus functions as a 'stop' or 'termination' codon, leading to termination of chain synthesis. If this type of mutation is near the 3' end of the gene, an abnormal but functional globin chain is produced; however, if it is more proximal, the chain produced is likely to be not only short but very unstable, leading to a dominant thalassaemia phenotype. Point mutations can also convert a stop codon to a coding sequence, so that an elongated mRNA and elongated globin chain are produced. A variant haemoglobin with two amino acid substitutions resulting from two point mutations can be produced either from a new mutation occurring in the gene encoding a variant globin chain, e.g. in a parental germ cell, or from cross-over between two variant alleles.

An unusual result of a point mutation is the production of an abnormal amino acid that is converted to a different amino acid by post-translational modification. This may be the result of deamidation, acetylation or oxidation. There are six reported variant haemoglobins in which the abnormal DNA sequence codes for asparagine, but this is subsequently deamidated to aspartic acid [11]; of these, the most common is haemoglobin J-Sardegna ($\alpha 50(CD8)^{His \rightarrow Asn \rightarrow Asp}$), which has a prevalence of 0.25% in northern Sardinia. Post-translational acetylation occurs in haemoglobin Raleigh, which has a $\beta 1^{Val \rightarrow Ala}$ substitution; proteins with an N-terminal alanine are often acetylated and this is the case with this variant haemoglobin [35]. The presence in one individual of haemoglobins with three different β chains may be attributable to post-translational modification. For example, the replacement of leucine by hydroxyleucine that characterizes haemoglobin Coventry is not encoded by genomic DNA and is found only in the presence of an unstable haemoglobin, either haemoglobin Atlanta or haemoglobin Sydney. Some mutations affecting the haem pocket and leading to haemoglobin instability permit the oxidation of leucine to isoleucine [36]. Haemoglobin Bristol also shows post-translational modification. It is an unstable haemoglobin resulting from conversion of the β67 valine codon to a codon for methionine; however, the final haemoglobin has aspartic acid rather than methionine as a consequence of post-translational modification [11].

In a slightly different mechanism, the abnormal structure of a variant haemoglobin resulting from a point mutation leads to post-translational modification of a normal amino acid, in three cases leucine being modified to hydroxyleucine [11] and in one case asparagine adjacent to the abnormal residue being deamidated to aspartic acid [35].

Mutations in the codon for the N-terminal valine may mean that a different amino acid is encoded, with resultant retention of the initiator methionine and full acetylation of the N-terminal residue (e.g. the glutamate of the α chain variant haemoglobin

HAD1 8/5/05 2:11 PM Page 19

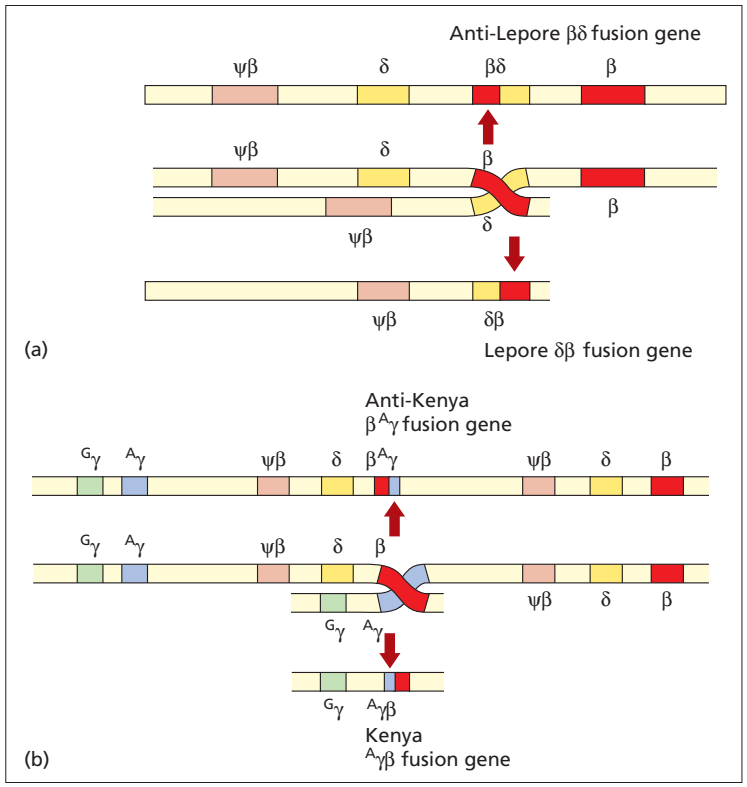

Fig. 1.14 Some examples of fusion genes produced by non-homologous cross-over: (a) formation of genes encoding Lepore and anti-Lepore haemoglobins; (b) formation of genes encoding Kenya and anti-Kenya haemoglobins.

Thionville) or normal cleavage of methionine but full acetylation of the N-terminal residue (e.g. the alanine of the α chain variant haemoglobin Lyon-Bron) [19]. Similarly, a histidine to proline change in position β2 leads to retention of the initiator methionine [35]. If methionine is retained, the globin chain is extended by one residue.

Deletions and insertions can lead to a frame shift, i.e. unless the deletion or insertion involves three nucleotides or multiples of three, the nucleotide sequences beyond the mutation will be in a different reading frame and will be 'read' during translation as coding for a completely different sequence of amino acids. Frame shift mutations can lead to a premature stop codon, so that both mRNA and the resultant globin chain are shorter than normal. Unless this occurs, a frame shift mutation is likely to lead to elongated mRNA and an elongated globin chain. The original stop codon is no longer in the reading frame and transcription continues until another stop codon is encountered.

Small deletions and large deletions and insertions can result from non-homologous cross-over between a pair of chromosomes during meiosis. These are usually in-frame. Non-homologous cross-over can involve not only a single pair of allelic genes (e.g. two α genes), but also two structurally similar but non-allelic genes (e.g. a β gene and a δ gene); in the latter instance, there may be a loss of the two normal genes and the production of a fusion gene which has 5′ sequences of one gene and 3′ sequences of the other gene; alternatively, the two normal genes may be retained with part of both genes being reduplicated in the fusion gene. Some examples of non-homologous cross-over are shown in Fig. 1.14. Non-homologous cross-over can also result in the reduplication of genes; for example, some individuals, instead of having two α genes on each chromosome 16, have three or even four α genes on one chromosome. Duplicated α genes occur in many populations and in some are quite frequent. For example, 2% of Sri Lankans have ααα.

HAD1 8/5/05 2:11 PM Page 20

Very rarely, individuals are somatic mosaics, so that a variant haemoglobin is present in an unusually low percentage. For example, a patient has been reported with haemoglobin Korle Bu as a minor fraction as a result of constitutional mosaicism [37].

Haemoglobin dimers are stable, but the tetramers that they form are able to dissociate and re-associate. When both normal and variant haemoglobins are present, heterotetramers and homotetramers will be found *in vivo*; for example, in the case of sickle cell trait, there will be $\alpha_2\beta_2$, $\alpha_2\beta^S{}_2$ and $\alpha_2\beta\beta^S$. When haemoglobins are studied *in vitro*, e.g. by electrophoresis or chromatography, the heterotetramers dissociate and re-associate as homotetramers. Some variant haemoglobins have abnormally stable tetramers, so that three rather than two forms are detected by haemoglobin electrophoresis and similar techniques.

Thalassaemias and haemoglobinopathies

Mutations can lead not only to the synthesis of a structurally abnormal haemoglobin, but also to a reduced rate of synthesis of a globin chain and therefore of the haemoglobin species of which it forms a part. The term 'thalassaemia' is used to describe disorders with a significant decrease in the rate of synthesis of one or more globin chains: α thalassaemia indicates a reduced rate of synthesis of α globin chain; similarly, β, δ, δβ and γδβ thalassaemias indicate a reduced rate of synthesis of β, δ, δ + β and γ + δ + β chains, respectively. In some disorders, there is both synthesis of a structurally abnormal haemoglobin and a reduced rate of synthesis of the variant haemoglobin. This is the case, for different reasons, with the α chain variant haemoglobin Constant Spring (first described in a Chinese patient in Constant Spring, a district of Kingston, Jamaica) and the β chain variant haemoglobin E. The term 'haemoglobinopathy' is sometimes used to indicate only those disorders with a structurally abnormal haemoglobin, while others use the term to include all disorders of globin chain synthesis, encompassing also the thalassaemias. If the term 'haemoglobinopathy' is used only to designate a structurally abnormal haemoglobin, variant haemoglobins, such as haemoglobin E and haemoglobin Constant Spring, can be referred to as 'thalassaemic haemoglobinopathies'.

Haemoglobinopathies may result from mutation of a β globin gene, in which case there is a variant form of haemoglobin A, or from mutation of an α globin gene, in which case there are variant forms of haemoglobins F, A and A_2. Similarly, mutations of γ and δ genes result in mutant forms of haemoglobin F and haemoglobin A_2, respectively. Because there are two β genes, an individual can have both a β chain variant and haemoglobin A or two β chain variants. Because there are usually four α genes, an individual could, in theory, have up to four different α chain variants; in practice, a number of individuals have been described with both haemoglobin A and two different α chain variants, e.g. haemoglobin Buda, haemoglobin Pest and haemoglobin A in one instance and haemoglobin G-Philadelphia, haemoglobin J-Sardegna and haemoglobin A in several instances.

The proportion of variant haemoglobins

The proportion of an α chain variant in the blood might be expected to be around 25%, as there are usually four α genes. However, the situation is far more complex. The variant is likely to be more than 25% if it results from mutation of the α2 gene (as the ratio of α2 to α1 synthesis is normally about 3 : 1) and less than 25% if it results from mutation of the α1 gene. The percentage is raised if there is coinheritance of α thalassaemia and lowered if there is coinheritance of triple α (ααα). If a gene encoding an α chain variant is a mutated α1 gene in *cis* with deletion of the α2 gene, it may be upregulated, increasing the percentage further. The percentage is reduced if the variant α chain is synthesized at a reduced rate, if it has a lower affinity for β chains than does the normal α chain or if the variant α chain or the variant haemoglobin is unstable.

Similarly, it might be expected that a β chain variant would represent about 50% of total haemoglobin in heterozygotes, as there are two β genes. As for α chain variants, the situation is much more complex. The percentage may be above 50% in the case of variants with negatively charged β chains, which have a greater affinity than normal β chains for the positively charged normal α chains (e.g. haemoglobin J-Baltimore or J-Iran); if there is coexisting α

HAD1 8/5/05 2:11 PM Page 21

Table 1.5 Consequences of mutation of globin genes.

Type of mutation and consequence	Example
Substitution of an external amino acid not involved in inter-chain contacts; no functional abnormality	Haemoglobin G-Philadelphia
Amino acid substitution leading to reduced solubility, polymerization of haemoglobin and deformation of cells into a holly leaf or sickle shape with consequent haemolysis and vascular obstruction	Haemoglobin S (sickle cell haemoglobin)
Amino acid substitution leading to reduced solubility, formation of straight-edged crystals and haemolysis	Haemoglobin C
Replacement of haem-binding or haem-related histidine residue by another amino acid leading to an increased tendency to oxidation, i.e. formation of methaemoglobin; there is cyanosis at birth if the defect is in a γ gene, cyanosis from birth if the defect is in an α chain and cyanosis from approximately 6 months of age if the defect is in a β chain; there may be associated haemoglobin instability	M haemoglobins
Mutation involving amino acids of the haem pocket or $\alpha_1\beta_2$ (tetrameric) contacts or mutation interfering with the helical structure of haemoglobin, leading to haemoglobin instability and Heinz body haemolytic anaemia; there may also be decreased oxygen affinity and consequent cyanosis	Haemoglobin Köln, haemoglobin Zurich (haem pocket mutation), haemoglobin Kansas (mutation affecting α1β2 contacts)
Mutations involving $\alpha_1\beta_2$, $\alpha_2\beta_1$ tetrameric haemoglobin contacts or C-terminal end of β chain, where there are residues involved in 2,3-DPG interaction and stability of the deoxy form of haemoglobin, leading to increased oxygen affinity and polycythaemia	Haemoglobin Chesapeake, haemoglobin Bethesda, haemoglobin Kempsey, haemoglobin J-Capetown, haemoglobin Yakima
Mutation leading to decreased oxygen affinity and therefore anaemia, as normal tissue delivery of oxygen is achieved with a lower concentration of haemoglobin; may cause cyanosis	Haemoglobin S, haemoglobin Seattle (also unstable), haemoglobin Kansas (also unstable), haemoglobin Beth Israel
Mutation in β gene leading to markedly reduced or absent β chain production, reduced synthesis of haemoglobin A and possibly ineffective erythropoiesis consequent on damage to developing erythroblasts by excess α chains	β thalassaemia (major, intermedia or minor)
Mutation in β gene leading to structurally abnormal and very unstable β chain	(Dominant) β thalassaemia phenotype
Mutation in α gene leading to markedly reduced or absent α chain synthesis and reduced synthesis of haemoglobins F, A and A_2	α thalassaemia (α thalassaemia trait, haemoglobin H disease or haemoglobin Bart's hydrops fetalis)
Mutation in α gene leading to structurally abnormal α chain synthesized at a greatly reduced rate	α thalassaemia phenotype, e.g. haemoglobin Constant Spring
Mutation in δ gene leading to a structural abnormality or markedly reduced or absent δ chain production	Haemoglobin A_2 variant or δ thalassaemia; no clinical significance as haemoglobin A_2 is a minor haemoglobin, but complicates the diagnosis of thalassaemia trait
Mutation in γ gene leading to structural abnormality or reduced rate of synthesis of γ chain and therefore haemoglobin F	Some methaemoglobins

2,3-DPG, 2,3-diphosphoglycerate.

HAD1 8/5/05 2:11 PM Page 22

thalassaemia, leading to a lack of α chains, the percentage of the variant is even higher. The converse is seen with positively charged β chains, such as β^S, β^C, $\beta^{O\text{-}Arab}$ and $\beta^{D\text{-}Punjab}$, which have a lower affinity than normal β chains for normal α chains. The percentage of the variant is thus somewhat less than 50% and, if there is coexisting α thalassaemia, is even lower. The percentage is also reduced considerably below 50% if there is a reduced rate of synthesis of the variant β (or δβ) chain (e.g. β^E, $\delta\beta^{Lepore}$), if the β chain is unstable or if the variant haemoglobin is unstable (e.g. haemoglobin Köln).

An alteration in the amino acid sequence of the globin chains, i.e. an alteration in the primary structure of haemoglobin, often has no significant effect on the secondary, tertiary and quaternary structure of haemoglobin; this is the case when the substituted amino acid is of a similar size to the normal amino acid, has the same charge and the same hydrophobic or hydrophilic properties, and does not have a role in the binding of haem or 2,3-DPG or in interactions between chains. In this case, a variant haemoglobin has no consequences for the health of the individual. In other cases, an alteration in the primary structure of haemoglobin affects the secondary, tertiary or quaternary structure of the molecule, sometimes with very profound effects. Some of the effects of mutations in globin genes are shown in Table 1.5.

Over 1000 mutations of the globin genes have been recognized. Some 690 of them have been collated in a single volume [38] and this database is now available electronically, in updated form, on the World Wide Web (http://globin.cse.psu.edu/).

Check your knowledge

One to five answers may be correct. Answers to almost all questions can be found in this chapter or can be deduced from the information given. The correct answers are given on p. 25.

1.1 The haemoglobin molecule
- (a) requires iron for its synthesis
- (b) is composed of three pairs of globin chains
- (c) alters its structure when oxygen is bound
- (d) is assembled in the cytosol
- (e) binds 2,3-diphosphoglycerate

1.2 Haemoglobin F
- (a) is the major haemoglobin present in the fetus
- (b) has a lower oxygen affinity than haemoglobin A
- (c) is absent in normal adults
- (d) percentage shows a non-Gaussian distribution in the population
- (e) is composed of two α chains and two β chains

1.3 The functions of haemoglobin include
- (a) transport of glucose
- (b) transport of CO_2
- (c) transport of oxygen
- (d) buffering
- (e) transport of creatinine to the kidney

1.4 The affinity of haemoglobin for oxygen is decreased by
- (a) fever
- (b) alkalosis
- (c) binding of CO_2
- (d) binding of 2,3-diphosphoglycerate
- (e) glycosylation

1.5 When blood circulates through the lungs, haemoglobin
- (a) is oxidized
- (b) takes up oxygen
- (c) loses CO_2
- (d) takes up water
- (e) dissociates into haem and globin

1.6 Structurally abnormal haemoglobins may result from
- (a) point mutations
- (b) gene fusion
- (c) frame shift mutations
- (d) mutation of a stop codon to a coding sequence
- (e) mutation of a coding sequence to a stop codon

1.7 Abnormal haemoglobins may
- (a) have increased oxygen affinity
- (b) have decreased oxygen affinity
- (c) be prone to crystallize

HAD1 8/5/05 2:11 PM Page 23

(d) be unstable
(e) be abnormally prone to oxidation

1.8 Mutations in globin genes
(a) can occur in α, β, $^{G}\gamma$, $^{A}\gamma$ and δ genes
(b) always result in a structural abnormality of haemoglobin
(c) always have harmful effects
(d) can lead to a reduced rate of globin chain synthesis
(e) can convert one gene to another

1.9 Haemoglobin F
(a) is present, in adult life, in a subset of erythrocytes referred to as F cells
(b) is composed of two α chains and two γ chains, encoded by two pairs of structurally similar α genes and two pairs of structurally similar γ genes
(c) has a sigmoid dissociation curve
(d) constitutes a higher proportion of total haemoglobin in premature than in full-term babies
(e) on average is present at a somewhat higher level in women than in men

1.10 Cooperativity is essential for
(a) a sigmoid oxygen dissociation curve
(b) the higher oxygen affinity of haemoglobin F in comparison with haemoglobin A
(c) the Bohr effect
(d) the binding of CO_2 to haemoglobin
(e) conversion of haemoglobin to methaemoglobin

1.11 The proportion of a variant haemoglobin is usually
(a) greater in the case of an α chain variant than a β chain variant
(b) greater in the case of an α chain variant if there is coexisting deletion of an α gene
(c) greater if the variant β chain has a higher affinity for normal α chains than does the normal β chain
(d) greater, in the case of haemoglobin S, if there is coexisting α thalassaemia
(e) greater if the variant haemoglobin is unstable

Further reading

Bunn HF and Forget BG. *Hemoglobin: Molecular, Genetic and Clinical Aspects*. W. B. Saunders, Philadelphia, PA, 1986.

Huisman THJ, Carver MFH and Efremov GD. *A Syllabus of Human Hemoglobin Variants*. The Sickle Cell Anemia Foundation, Augusta, GA, 1996.

Lehmann H and Huntsman RG. *Man's Haemoglobins including the Haemoglobinopathies and their Investigation*. North Holland Publishing Company, Amsterdam, 1974.

Nienhuis AW and Benz EJ (1997) Regulation of hemoglobin synthesis during the development of the red cell. *N Engl J Med* **297**, 1318–1328, 1371–1381, 1430–1436.

Stamatoyannopoulos G and Nienhuis AW. Hemoglobin switching. In: Stamatoyannopoulos G, Nienhuis AW, Majerus PW and Varmus H, eds. *The Molecular Basis of Blood Diseases*, 2nd edn. W. B. Saunders, Philadelphia, PA, 1994, pp. 107–156.

Steinberg MH, Forget BG, Higgs DR and Nagel RL, eds. *Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management*. Cambridge University Press, Cambridge, 2001.

The Globin Gene Server, hosted by Pennsylvania State University, USA and McMaster University, Canada. URL http://globin.cse.psu.edu/.

Weatherall D and Clegg JB. *The Thalassaemia Syndromes*, 4th edn. Blackwell Science, Oxford, 2001.

References

1 Gross GS and Lane P (1999) Physiological reactions of nitric oxide and hemoglobin: a radical rethink. *Proc Natl Acad Sci USA* **96**, 9967–9969.

2 Nagel RL and Jaffé ER. CO-, NO-, met-, and sulfhemoglobinemias: the dyshemoglobinemias. In: Steinberg MH, Forget BG, Higgs DR and Nagel RL, eds. *Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management*. Cambridge University Press, Cambridge, 2001, pp. 1214–1233.

3 Schechter AN and Gladwin MT (2003) Hemoglobin and the paracrine and endocrine functions of nitric oxide. *N Engl J Med* **348**, 1483–1485.

4 Pauling L, Itano HA, Singer SY and Wells IG (1949) Sickle-cell anemia, a molecular disease. *Science* **110**, 543–548.

5 Kunkel HG and Wallenius G (1955) New hemoglobin in normal adult blood. *Science* **122**, 288.

HAD1 8/5/05 2:11 PM Page 24

6 Bunn HF and Forget BG. *Hemoglobin: Molecular, Genetic and Clinical Aspects*. W. B. Saunders, Philadelphia, PA, 1986.

7 Huehns ER, Flynn FV, Butler EA and Beaven GH (1961) Two new hemoglobin variants in the very young human embryo. *Nature* **189**, 496–497.

8 Hecht F, Jones RT and Koler RD (1967) Newborn infants with Hb Portland 1, an indicator of α-chain deficiency. *Ann Hum Genet* **31**, 215–218.

9 Fantoni A, Farace MG and Gambari R (1981) Embryonic hemoglobins in man and other mammals. *Blood* **57**, 623–633.

10 Nagel RL and Steinberg MH. Hemoglobins of the embryo and fetus and minor hemoglobins of adults. In: Steinberg MH, Forget BG, Higgs DR and Nagel RL, eds. *Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management*. Cambridge University Press, Cambridge, 2001, pp. 197–230.

11 Rees DC, Rochette J, Schofield C, Green B, Morris M, Parker NE *et al*. (1996) A novel silent posttranslational mechanism converts methionine to aspartate in hemoglobin Bristol (β67[E11] Val-Met→Asp). *Blood* **88**, 341–348.

12 Ponka P (1997) Tissue-specific regulation of iron metabolism and heme synthesis: distinct control mechanisms in erythroid cells. *Blood* **89**, 1–25.

13 Steinberg MH and Adams JG (1991) Haemoglobin A_2: origin, evolution, and aftermath. *Blood* **78**, 2165–2177.

14 Stamatoyannopoulos G and Nienhuis AW. Hemoglobin switching. In: Stamatoyannopoulos G, Nienhuis AW, Majerus PW and Varmus H, eds. *The Molecular Basis of Blood Diseases*, 2nd edn. W. B. Saunders, Philadelphia, PA, 1994, pp. 107–156.

15 Weatherall D and Clegg JB. *The Thalassaemia Syndromes*, 4th edn. Blackwell Science, Oxford, 2001.

16 Ho PJ and Thein SL (2000) Gene regulation and deregulation: a β globin perspective. *Blood Rev* **14**, 78–93.

17 Russell JE and Liebhaber SA (1996) The stability of human β-globin mRNA is dependent on structural determinants positioned within its 3′ untranslated region. *Blood* **87**, 5314–5322.

18 Jane SM and Cunningham JM (1998) Understanding fetal globin gene expression: a step towards effective Hb F reactivation in haemoglobinopathies. *Br J Haematol* **102**, 415–422.

19 Lacan P, Souillet G, Aubry M, Prome D, Richelme-David S, Kister J *et al*. (2002) New alpha 2 globin chain variant with low oxygen affinity affecting the N-terminal residue and leading to N-acetylation [Hb Lyon-Bron α1(NA1)Val→Ac-Ala]. *Am J Hematol* **69**, 214–218.

20 Bernini LF and Harteveld CL (1998) α-Thalassaemia. *Bailliere's Clin Haematol* **11**, 53–90.

21 Serjeant BE, Mason KP and Serjeant GR (1978) The development of haemoglobin A_2 in normal Negro infants and in sickle cell disease. *Br J Haematol* **39**, 259–263.

22 Steinberg MH (1998) Pathophysiology of sickle cell disease. *Bailliere's Clin Haematol* **11**, 163–184.

23 Huehns ER and Beaven GH. Developmental changes in human haemoglobins. In: Benson PF, ed. *The Biochemistry of Human Development*. Spastics International Medical Publications, London, 1971, p. 175. Quoted by Weatherall D and Clegg JB. *The Thalassaemia Syndromes*, 3rd edn. Blackwell Scientific Publications, Oxford, 1981.

24 Colombo B, Kim B, Atencio RP, Molina C and Terrenato L (1976) The pattern of foetal haemoglobin disappearance after birth. *Br J Haematol* **32**, 79–87.

25 Cheron G, Bachoux I, Maier M, Massonneau M, Peltier JY and Girot T (1989) Fetal hemoglobin in sudden infant death syndrome. *N Engl J Med* **320**, 1011–1012.

26 Garner C, Tatu T, Reittie JE, Littlewood T, Darley J, Cervino S *et al*. (2000) Genetic influences on F cells and other hematologic variables: a twin heritability study. *Blood* **95**, 342–346.

27 Miyoshi K, Kaneto Y, Kawai H, Ohchi H, Niki S, Hasegawa K *et al*. (1988) X-linked dominant control of F-cells in normal adult life: characterization of the Swiss type as hereditary persistence of fetal hemoglobin regulated dominantly by genes on X chromosome. *Blood* **72**, 1854–1860.

28 Merghoub T, Perichon B, Maier-Redelsperger M, Dibenedetto SP, Samperi P, Ducrocq R *et al*. (1997) Dissection of the association status of two polymorphisms in the β-globin gene cluster with variations in F-cell number in non-anemic individuals. *Am J Hematol* **56**, 239–243.

29 Craig JE, Rochette J, Sampietro M, Wilkie AOM, Barnetson R, Hatton CSR *et al*. (1997) Genetic heterogeneity in heterocellular hereditary persistence of fetal hemoglobin. *Blood* **90**, 428–434.

30 Tadmouri GO, Yuksel L and Basak AN (1998) Hb S/beta (del) thalassemia associated with high hemoglobin A_2 and F in a Turkish family. *Am J Hematol* **59**, 83–86.

HAD1 8/5/05 2:11 PM Page 25

31 Zertal-Zidani S, Ducrocq R, Weil-Oliver C, Elion J and Krishnamoorthy R (2001) A novel δβ fusion gene expresses hemoglobin A (HbA) not Hb Lepore: Senegalese $\delta^0\beta^+$ thalassaemia. *Blood* **98**, 1261–1263.
32 Viprakasit V, Gibbons RJ, Broughton BC, Tolmie JL, Brown D, Lunt P *et al*. (2001) Mutations in the general transcription factor TFIIH result in beta-thalassaemia in individuals with trichothiodystrophy. *Hum Mol Genet* **10**, 2797–2802.
33 Yu C, Niakan KK, Matsushita M, Stamatoyannopoulos G, Orkin SH and Raskind WH (2002) X-Linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. *Blood* **100**, 2040–2045.
34 Huisman THJ (1997) Gamma chain abnormal human fetal hemoglobin variants. *Am J Hematol* **55**, 159–163.
35 Steinberg MH and Nagel RL. New and recombinant mutant hemoglobins of biological interest. In: Steinberg MH, Forget BG, Higgs DR and Nagel RL, eds. *Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management*. Cambridge University Press, Cambridge, 2001, pp. 1195–1211.
36 Brennan SO, Shaw J, Allen J and George PM (1992) β141 Leu is not deleted in the unstable haemoglobin Atlanta-Coventry but is replaced by a novel amino acid of mass 128 daltons. *Br J Haematol* **81**, 99–103.
37 Wild BJ, Green BN, Lalloz MRA and Layton DM (2000) When is a minor haemoglobin fraction worthy of investigation? *Br J Haematol* **108** (Suppl. 1), 40.
38 Huisman THJ, Carver MFH and Efremov GD. *A Syllabus of Human Hemoglobin Variants*. The Sickle Cell Anemia Foundation, Augusta, GA, 1996.

Answers to questions

1.1	(a) T	1.3	(a) F	1.5	(a) F	1.7	(a) T	1.9	(a) T	1.11	(a) F
	(b) F		(b) T		(b) T		(b) T		(b) T		(b) T
	(c) T		(c) T		(c) T		(c) T		(c) T		(c) T
	(d) T		(d) T		(d) F		(d) T		(d) T		(d) F
	(e) T		(e) F		(e) F		(e) T		(e) T		(e) F
1.2	(a) T	1.4	(a) T	1.6	(a) T	1.8	(a) T	1.10	(a) T		
	(b) F		(b) F		(b) T		(b) F		(b) F		
	(c) F		(c) T		(c) T		(c) F		(c) T		
	(d) T		(d) T		(d) T		(d) T		(d) F		
	(e) F		(e) F		(e) T		(e) T		(e) F		