MEDICAL PRACTICE

Occasional Survey

Molecular Basis for Some Disorders of Haemoglobin Synthesis*—I

D. J. WEATHERALL

British Medical Journal, 1974, 2, 451-454**Summary**

The genetic disorders of haemoglobin produce a wide range of clinical syndromes which pose a major public health problem in many parts of the world. Investigation of these conditions at the molecular level has uncovered an extraordinary diversity of mutational events which can produce human disease. These include point mutations, loss of one or more base triplets of the DNA of the haemoglobin genes, insertion of new genetic material, the production of fusion genes due to abnormal crossing over, frame shifts, chain-termination mutations, and the deletion (loss) of complete genes. It is possible in many cases to relate these molecular abnormalities to the clinical pictures which they produce.

Introduction

Much has been learnt about the molecular pathology of the abnormal haemoglobin disorders since the chance conversation between William Castle and Linus Pauling in 1945 which led to the discovery that sickle-cell anaemia is due to a structural alteration in the haemoglobin molecule.^{1 2} Indeed more is known about the molecular mechanisms which underlie these conditions than any other group of human diseases. These advances

have been particularly important because the haemoglobinopathies are probably the commonest clinical disorders to result from single gene mutations and they produce a vast public health problem in many parts of the world. Furthermore, lessons learnt from the abnormal haemoglobin field have a wide application to our understanding of many other genetic and acquired disorders.

Despite the current view that molecular biology is an exciting intellectual exercise of little practical value, expressed so eloquently by Burnet,³ many of the recent advances in the abnormal haemoglobin field have followed the development of new techniques in this much maligned discipline. I shall review these advances and speculate briefly about their possible applications to the prevention and treatment of the haemoglobinopathies.

Normal Human Haemoglobin

To appreciate recent events in the field of abnormal haemoglobin it is necessary to have a working acquaintance with the structure and synthesis of normal haemoglobin. This has been described in detail in several recent reviews⁴⁻⁶ and will be outlined only briefly here.

STRUCTURE

Human haemoglobin is a globular protein which consists of four peptide chains. Normal adult haemoglobin (A) has two α -chains made up of 141 amino-acids and two β -chains which are 146 amino-acids long; it is written $\alpha_2\beta_2$. A haem molecule, a complex protoporphyrin-ring structure containing iron, is attached to each of these chains. In fetal life the main respiratory protein is haemoglobin F which has α -chains identical to those of haemoglobin A but in this case paired with γ -chains; it is written $\alpha_2\gamma_2$.

*Based on the 1974 Watson Smith lecture given at the Royal College of Physicians of London on 8 May 1974.

Department of Haematology, University of Liverpool, Liverpool L69 3BX

D. J. WEATHERALL, F.R.C.P., M.R.C.PATH., Professor of Haematology.
(Present address: Nuffield Department of Clinical Medicine Radcliffe Infirmary, Oxford OX2 6HE)

Because of the interactions of the side chains of individual amino-acids which make up the globin chains each folds into a coiled three-dimensional structure which makes a compact fit with each of its fellow chains to form the complex structure which haemoglobin presents to the x-ray crystallographer (fig. 1). The accurate reproduction of this three-dimensional shape is vital if the molecule is to perform efficiently as an oxygen carrier.

FIG. 1—Model of α - and β -chains of human haemoglobin. Complete molecule is formed by superimposing α -chains on β -chains by moving them in direction of arrow. Contact areas are shown by continuous dark and light lines on α - and β -chains respectively. HS = Reactive sulphhydryl group. N = Amino-terminal end of peptide chain. (Reproduced by permission of M. F. Perutz.)

The important functional areas of the haemoglobin molecule are the contacts between the globin chains and the deep pockets in their surfaces which house the haem groups. Each α -chain is in contact with both β -chains at critical "hinge" areas called the $\alpha_1\beta_1$ and $\alpha_1\beta_2$ contacts, but they make relatively weak chemical contacts with their fellow α -chains. Similarly the β -chains while making strong chemical contacts at the $\alpha\beta$ hinges make only weak bonds with each other. It has been realized for many years that the production of the sigmoid oxygen dissociation curve of haemoglobin, without which, as Barcroft pointed out, man would be able to perform little better than the lobster, requires interaction between the four haem groups. The molecular basis for this important reaction is now understood, at least in outline.⁷⁻⁹ During oxygenation and deoxygenation a complex series of spacial alterations occur both within and between the haemoglobin subunits, which facilitate the uptake or release of oxygen after the first haem group has been oxygenated or deoxygenated. Furthermore, for the iron of the haem to carry out the remarkable function of reversible oxygenation it must be protected in a water-repellent pocket. Indeed the structure of the latter must be critical since it has changed very little during evolution. It follows, therefore, that for the normal function and stability of the molecule its correct three dimensional configuration is critical, and thus even a single amino-acid substitution can alter both properties profoundly.

SYNTHESIS AND GENETIC CONTROL

The genetic information which determines the order of amino-acids in the globin chains is contained in the DNA of their respective genes by virtue of the order of the constituent bases, adenine (A), thymine (T), guanine (G), and cytosine (C). Because of the rules of base pairing A always pairs with T and G with C in the DNA helix. The genetic code is a three-letter (base) code—that is, three bases code for a particular amino-acid. Furthermore the code is "degenerate," meaning that there is more than one code word for each amino-acid. The information stored in the DNA has to be transported from the nucleus to the

cytoplasm and this is achieved by an intermediary, a form of ribonucleic acid called messenger RNA (mRNA). When an mRNA molecule is to be transcribed the DNA helix opens and one strand is copied. RNA is very similar in structure to DNA except that it has a different sugar, deoxyribose instead of ribose, and the base thymine (T) is replaced by uracil (U) which can only form a base pair with adenine. Thus when a DNA strand is copied, a process which is catalysed by an enzyme called RNA polymerase, the resulting mRNA strand because of the rules of base pairing makes a faithful copy of its DNA template. It is now clear that mRNA is made initially as a large precursor molecule, called heterogeneous nuclear RNA (HnRNA), which is processed to produce a smaller molecule which diffuses into the cell cytoplasm and acts as a template for protein synthesis.

Protein synthesis occurs in three stages called initiation, elongation, and termination. Amino-acids which are to make up a globin chain cannot interact directly with mRNA and have to be transported to the correct position by "carrier" molecules called transfer RNAs (tRNAs). Each tRNA has three bases, called an anticodon, which are complementary to appropriate triplets of the mRNA. Because of the degeneracy of the genetic code it follows that most amino-acids have several different species of tRNA. To initiate protein synthesis a particular tRNA which carries methionine is required. There is a specific "start word" (base triplet) in the mRNA, which is probably AUG, and the initiating tRNA binds to this site. The growing peptide chain is carried, as it is elongated, by the ribosomes which are spherical macromolecules made up of RNA and protein. The ribosomes have a large and small subunit which together act as a kind of pulley which holds the growing chain and provide a series of sites for tRNAs and enzymes involved in chain elongation. Initiation is completed by the association of the large and small ribosomal subunits together with the initiating transfer RNA and a complex series of initiating factors. Once the first amino-acid is in position the second amino-acid moves into place attached to its tRNA and a peptide bond is formed between the two; the globin chain is now two amino-acids (or residues) long. This procedure continues as the ribosomes move over the mRNA and the chain grows in length. The processes of elongation and translation continue until a specific base triplet reading "end" chain is arrived at, whereupon the ribosomes drop off the mRNA and the completed peptide chain is released. The haem part of the molecule is synthesized separately by a series of enzymatically controlled steps, starting with active succinate from the Krebs cycle and culminating in the complex protoporphyrin ring with iron in the centre. As the globin chains are synthesized they fold up and haem is inserted into its pocket, and once they are released the completed chains combine to form a haemoglobin tetramer. The major steps in human haemoglobin synthesis and its genetic control are summarized in figs. 2 and 3.

FIG. 2—Genetic control of protein synthesis (see text).

FIG. 3—Genetic control of human fetal and adult haemoglobin production (see text).

The way in which this beautifully co-ordinated system is controlled is not yet clear but most of the available evidence points to mechanisms which control the rate of transcription of mRNA as being the most important factors in controlling the final level of haemoglobin in cells. The synthesis of human haemoglobin has been reviewed in detail by Clegg⁸ and Nienhuis and Anderson.¹⁰

Abnormal Haemoglobin Disorders

The genetic disorders of haemoglobin synthesis can be divided broadly into two main groups. Firstly, those in which a structural alteration in a globin chain alters the stability of function of the molecule. Secondly, conditions which result from a defective rate of synthesis of one of the globin chains, and in these disorders, known collectively as the thalassaemia syndromes, the structure of the globin chain may or may not be normal.

In many populations both types of genetic defect are common and therefore it is not at all rare to see individuals with both structural haemoglobin variants and thalassaemia, and so there are many genetic disorders of bewildering complexity which make up the whole spectrum of the haemoglobinopathies.

Structural Haemoglobin Variants

Over 120 human haemoglobin variants have now been characterized and in many cases it has been possible to relate the structural

change to abnormal function.¹¹ The molecular basis for these structural alterations is summarized briefly here.

SINGLE AMINO-ACID SUBSTITUTIONS

The simplest mechanism for the production of a haemoglobin variant is the substitution of a single base in the genetic code—that is, a point mutation. The genetic code was derived originally from work in an *Escherichia coli* system but it has been possible to show that all the amino-acid substitutions in the human haemoglobins reported to date are consistent with a single base change in the code as derived from micro-organisms (fig. 4). This extraordinary observation provides strong evidence that the code is universal throughout living organisms. It is sobering to reflect that the lowly *E. coli* codes for its proteins with the same codewords as the Nobel prizewinners who “broke” the genetic code.

Haemoglobin A β-chain				
Gene	TGA	GGT	CTC	CTC
mRNA	ACU	CCA	GAG	GAG
Globin	Thr	Pro	Glu	Glu

Haemoglobin S β-chain				
Gene	TGA	GGT	CAC	CTC
mRNA	ACU	CCA	GUG	GAG
Globin	Thr	Pro	Val	Glu
Position on chain	4	5	6	7

FIG. 4—Production of sickle-cell haemoglobin by single base change in the genetic code.

The result of a particular amino-acid substitution depends on the type of amino-acid that is changed and its position in the molecule. While many substitutions cause no abnormality there is a wide spectrum of clinical conditions which result from mutations which alter either the function or the stability of the molecule (or indeed both in some cases) (see table).

Shortening of red-cell survival due to single amino-acid substitutions may occur in several ways. The best known is the change of the molecular configuration which follows the substitution of valine for glutamic acid in position 6 in the β-chain, the sickling mutation.¹² This causes the haemoglobin molecules to become stacked up during deoxygenation though the exact mechanism whereby the valine substitution produces this effect is still a matter of conjecture.^{11 13} The red-cell survival may also

Molecular Pathology of Human Haemoglobin Variants, selected because each Variant is Example of One of a Particular Group of Variants which Produces a Clinical Disorder owing to a Specific Alteration in Configuration or Rate of Synthesis of Haemoglobin Molecules^{11 12 13}

Clinical Disorder	Molecular Change	Mechanism
Haemolytic Anaemia		
Sickle-cell Anaemia	β6 Glu → Val	Molecular stacking, red-cell distortion
Hb C Disease	β6 Glu → Lys	? Reduced solubility, rigid cell
Unstable Haemoglobins (Heinz-body anaemia):		
Hb Koln	β98 Val → Met	Change of conformation at critical bend or corner of helices, instability.
Hb Hamersmith	β42 Phe → Ser	Alters shape of haem pocket, instability.
Hb Philly	β35 Tyr → Phe	Weakens α ₁ β ₁ contact, instability.
Hb Gun Hill	β91-97 deleted	Removes haem and subunit contacts, instability.
Hb Genova	β28 Leu → Pro	Disrupts B helix, instability.
Drug-induced Haemolysis		
Hb Zurich	β63 His → Arg	Long side chain pushed to surface, instability only when exposed to oxidants.
Congenital Cyanosis		
Hb M Boston	β58 His → Tyr	Phenol group of Tyr forms bond with iron of haem, holding in stable ferric state—permanent methaemoglobinemia.
Hb M Iwate	β87 His → Tyr	Abnormal contact with haem ring, reduced oxygen affinity, instability
Hb Kansas	β87 Asn → Thr	
Congenital Polycythemia		
Hb Chesapeake	α92 Arg → Leu	Abnormal α ₁ β ₂ contact, high oxygen affinity.
Hb Little Rock	β143 His → Gln	Loss of 2,3 diphosphoglycerate binding, high oxygen affinity.
Hb Rainier	β145 Tyr → Cys	Alters conformational changes during oxygenation, high oxygen affinity.
Hypochromic Anaemia and Thalassaemia		
Hb Lepore	β8 fusion chain	Slowly synthesized, chain imbalance.
Hb Constant Spring	Chain termination → Glu	Slowly synthesized, chain imbalance.
Hb Icaria	Chain termination → Lys	Slowly synthesized, ? chain imbalance.
Hb K Woolwich	β132 Lys → Gln	? Slowly synthesized.

be shortened if the solubility of haemoglobin is reduced so that cells become more rigid than normal and cannot pass easily through the microvasculature. This is the basis of the mild haemolytic anaemia seen in the haemoglobin-C disorders.¹⁴ Finally, a whole series substitutions may alter the stability of the haemoglobin molecule. These conditions form the clinical syndrome of the unstable-haemoglobin disorders.^{15 16} Affected patients have a variable haemolytic anaemia with inclusion bodies in their red cells which result from instability of the affected haemoglobin molecules and their precipitation in the cell. In addition they may pass dark urine owing to the production of haem-breakdown products (dipyrroles) from the precipitated haemoglobin. The different types of substitution which can produce this clinical picture are summarized in the table.

Single amino-acid substitutions may also alter the function of haemoglobin. In some cases replacements near the haem pocket result in permanent methaemoglobin formation. There are a series of haemoglobins, all designated "M" followed by their place of discovery—for example Hb M Boston, M Saskatoon, which are associated with permanent cyanosis.¹⁷ In addition there is a steadily growing group of variants which are associated with hereditary polycythaemia.^{18 19} In these cases the amino-acid substitution results in an abnormal oxygen-dissociation curve and a high oxygen-affinity state such that the red cells give up relatively less oxygen in the tissues, which results in increased erythropoietin production and in an increased red-cell mass.

OTHER TYPES OF MUTATION CAUSING STRUCTURAL HAEMOGLOBIN VARIANTS

Some unstable haemoglobin variants result from a deletion (loss) of one or several amino-acids in a globin chain. The molecular mechanism for this unusual condition is probably chromosomal misalignment during meiosis with the loss of one or more base triplets due to abnormal crossing over. Usually the deletion results in an unstable molecule and hence in congenital inclusion-body anaemia.¹⁶ Other forms of genetic accidents can occur during meiosis. Unequal crossing over may cause additional genetic material to be inserted into a haemoglobin gene.^{20 21} If this caused lengthening of a globin chain at a critical area it could cause an unstable-haemoglobin disorder. In some cases fusion genes are formed, such as the $\delta\beta$ fusion gene which leads

to the production of Hb Lepore, a variant with normal α -chains combined with non- α -chains made up of the N-terminal residues of the δ -chain fused to the C-terminal residues of the β -chain.²² The composite $\delta\beta$ -chain is synthesized slowly and results in the clinical picture of thalassaemia. Indeed the Hb-Lepore abnormality was the first form of thalassaemia to be worked out at the molecular level. At least four fusion haemoglobin variants are now fully characterized. The latest, Hb Kenya, has non- α -chains which consist of part γ - and part β -chain sequences—that is, a composite fetal/adult globin chain.^{23 24}

The concluding part of this lecture will appear in next week's issue.

References

- Pauling, L., *et al.*, *Science*, 1949, 110, 543.
- Pauling, L., *The Harvey Lectures*, p. 216. New York, Academic Press, 1954.
- Burnet, M., *Genes, Dreams, and Realities*. Aylesbury, Medical and Technical Publishing, 1971.
- Huehns, E. R., in *Blood and Its Disorders*, ed. R. M. Hardisty and D. J. Weatherall. Oxford, Blackwell Scientific Publications. In press.
- Lehmann, H., and Huntsman, R. C., in *Clinics in Haematology*, ed. D. J. Weatherall. London, Saunders, In press.
- Clegg, J. B., in *Clinics in Haematology*, ed. D. J. Weatherall. London, Saunders. In press.
- Perutz, M. F., *et al.*, *Nature*, 1968, 219, 29.
- Perutz, M. F., *et al.*, *Nature*, 1968, 219, 131.
- Perutz, M. F., *Nature*, 1970, 228, 726.
- Nienhuis, A. W., and Anderson, W. F., in *Clinics in Haematology*, ed. D. J. Weatherall. London, Saunders. In press.
- Perutz, M. F., and Lehmann, H., *Nature*, 1968, 219, 902.
- Ingram, V. M., *Nature*, 1956, 178, 792.
- Milner, P. F., in *Clinics in Haematology*, ed. D. J. Weatherall. London, Saunders. In press.
- Charache, S., *et al.*, *Journal of Clinical Investigation*, 1967, 46, 1795.
- Grimes, A. J., Meisler, A., and Dacie, J. V., *British Journal of Haematology*, 1964, 10, 281.
- White, J. M., and Dacie, J. V., in *Progress in Haematology*, ed. E. B. Brown and C. V. Moore, Vol. 7, p. 69. New York, Grune and Stratton, 1971.
- Gerald, P. S., *Blood*, 1958, 13, 936.
- Charache, S., Weatherall, D. J., and Clegg, J. B., *Journal of Clinical Investigation*, 1966, 45, 813.
- Charache, S., in *Clinics in Haematology*, ed. D. J. Weatherall. London, Saunders. In press.
- Huisman, T. H. J., *et al.*, *American Society for Clinical Investigation, Abstracts*, p. 6, 1974.
- Charache, S., and Clegg, J. B. Personal communication, 1974.
- Baglioni, C., *Proceedings of the National Academy of Sciences of the United States of America*, 1962, 48, 1880.
- Huisman, T. H. J., *et al.*, *Archives of Biochemistry and Biophysics*, 1972, 15, 850.
- Smith, D. H., *et al.*, *Nature New Biology*, 1973, 246, 184.

Imported Diseases

Undiagnosed Fever—Rickettsial, Viral, and Helminth Infections

A. M. GEDDES

British Medical Journal, 1974, 4, 454-456

Rickettsial Infections

TYPHUS FEVER

Typhus fever is rarely imported into the United Kingdom. A recent traveller from Bangladesh developed headache, rigors, and fever six days after arriving in England. A provisional

diagnosis of malaria was excluded by examination of blood films and treatment was started with co-trimoxazole for suspected typhoid fever but without response. Cultures for *S. typhi* and the Widal reaction were negative. Re-examination of the patient showed an "eschar" on his back, and the Weil-Felix reaction subsequently confirmed the diagnosis of scrub typhus. This infection has an incubation period of 4-10 days and is characterized by headache, persistent fever, lymphadenopathy, splenomegaly, and a maculopapular rash. The eschar, which appears as a scab or ulcer, indicates the site of inoculation and multiplication of the infecting organism of scrub typhus, *Rickettsia orientalis*.

East Birmingham Hospital, Birmingham B9 5ST
A. M. GEDDES, M.B., F.R.C.P.ED., Consultant Physician