larly after the decline of the elms (*Ulmus*) (*25*), would similarly account for the more variable magnetic record.

Climatic Shifts

On a longer time scale, the magnetic mineral types and concentrations reflect the weathering, pedogenic, and denudational regimes prevailing through the major climatic shifts of the Pleistocene. Magnetic measurements can provide rapid insight into the nature of these shifts, their effect of weathering, soil development, and erosion rates, and their expression in the sedimentary record. Preliminary studies confirm paleoclimatic linkages in environments as diverse as northwest England (*26*), tropical Africa, and northern Queensland (*27*). The time scales of variation differ by an order of magnitude and the climatic regimes reflected span the range from arctic-alpine to humid tropical.

Prospect

Laboratory magnetic analyses have been shown to have many applications to a wide range of disciplines associated with environmental problems. The importance of these magnetic techniques will probably be further extended by the development of sensitive instruments capable of measuring, in the field, both magnetic remanence and magnetic susceptibility.

References and Notes

1. M. Faraday, *Philos. Trans. R. Soc. London* (1832), pp. 125-162.
2. D. W. Collinson, *Rev. Geophys. Space Phys.* **13**, 659 (1975).
3. F. Oldfield, R. Thompson, K. E. Barber, *Science* **199**, 679 (1978).
4. F. Oldfield, P. G. Appleby, R. S. Cambray, J. D. Eakins, K. E. Barber, R. W. Battarbee, G. R. Pearson, J. M. Williams, *Oikos* **33**, 40 (1979).
5. P. Pakarinen and K. Tolonen, *Suo* **28** (pt. 1), 19 (1977).
6. R. Thompson and F. Oldfield, *Phys. Earth Planet. Inter.* **17**, 300 (1978).
7. E. Le Borgne, *Ann. Geophys.* **11**, 399 (1955).
8. C. E. Mullins, *J. Soil Sci.* **28**, 223 (1977).
9. G. Longworth, L. W. Becker, R. Thompson, F. Oldfield, J. A. Dearing, T. A. Rummery, *ibid.* **30**, 93 (1979).
10. T. A. Rummery, J. Bloemendal, J. Dearing, F. Oldfield, R. Thompson, *Ann. Geophys.* **35**, 103 (1979).
11. G. Longworth and M. S. Tite, *Archaeometry* **19**, 3 (1977).
12. F. Oldfield, *Prog. Phys. Geogr.* **1**, 460 (1977).
13. _____, T. A. Rummery, R. Thompson, D. E. Walling, *Water Resour. Res.* **15**, 211 (1979).
14. D. Ninkovitch, N. D. Opdyke, B. C. Heezen, J. H. Foster, *Earth Planet. Sci. Lett.* **1**, 476 (1966).
15. N. D. Opdyke, *Rev. Geophys. Space Phys.* **10**, 213 (1972).
16. F. J. H. Mackereth, *Earth Planet. Sci. Lett.* **12**, 332 (1971).
17. L. A. Bauer, *Phys. Rev.* **3**, 34 (1899).
18. S. R. C. Malin and E. C. Bullard, in preparation.
19. M. J. Aitken, *Philos. Trans. R. Soc. London Ser. A* **269**, 77 (1970).
20. L. Molyneux, R. Thompson, F. Oldfield, M. E. McCallan, *Nature (London)* **237**, 4 (1972).
21. J. Bloemendal, F. Oldfield, R. Thompson, *ibid.* **280**, 50 (1979).
22. R. Thompson and D. J. Morton, *J. Sediment. Petrol.* **49**, 801 (1979).
23. R. Thompson, R. W. Battarbee, P. E. O'Sullivan, F. Oldfield, *Limnol. Oceanogr.* **20**, 687 (1975).
24. F. J. H. Mackereth, *Philos. Trans. R. Soc. London Ser. B* **250**, 165 (1966).
25. H. Godwin, *History of the British Flora* (Cambridge Univ. Press, Cambridge, ed. 2, 1975).
26. F. Oldfield, J. Dearing, R. Thompson, S. E. Garret-Jones, *Pol. Arch. Hydrobiol.* **25**, 321 (1978).
27. A. P. Kirshaw, *Nature (London)* **251**, 222 (1974).

Disorders of Human Hemoglobin

Arthur Bank, J. Gregory Mears, Francesco Ramirez

The human hemoglobin system is a model for the study of the regulation of specific eukaryotic genes (*1*). The complete structure of the major normal human globins is known, and a variety of mutant hemoglobins have led to extensive genetic and structural analysis. A group of anemias, the thalassemia syndromes, provide a series of mutants in which the biosynthesis of either α or β globin is decreased or absent. At least some of these anemias appear to be due to defects in the regulation of structurally normal globin genes. Restriction endonuclease analysis of cellular DNA and the cloning and sequencing of human globin genes have permitted detailed analysis of the organization and structure of normal and abnormal human globin genes. These techniques have provided a linear physical map of nucleotide sequences showing the organization of these genes. Deletions of specific nucleotide sequences, both within and surrounding the human globin genes, occur in certain anemias, and regions of DNA which may be important in regulating globin gene expression have been identified. Direct analysis can now be made of the relation between changes in gene structure to alterations in gene function; also prenatal diagnosis of certain disorders of man is now possible by analysis of deletions and base substitutions in fetal DNA from cells from amniotic fluid. In this article, we review normal and abnormal human hemoglobin synthesis, and focus on the insights provided by recent studies on the regulation of the human globin genes at the molecular level.

Dr. Bank is professor of medicine, and of human genetics and development, at Columbia University College of Physicians and Surgeons, New York 10032. Drs. Mears and Ramirez were formerly research associates at Columbia University. Dr. Mears is now an assistant professor of medicine at Albert Einstein College of Medicine, Bronx, New York 10461. Dr. Ramirez is now an assistant professor in the Department of Obstetrics and Gynecology at Rutgers University Medical School, Piscataway, New Jersey 08854.

Normal Hemoglobin Biosynthesis

The major hemoglobin genes and their protein products are shown in Fig. 1. There is strong genetic and biochemical evidence for linkage of four genes other than the α genes: two γ genes, one δ gene, and one β gene on chromosome 11 (*2*). The α genes are on chromosome 16, and appear to be duplicated (*3*). Two types of γ globin genes are identified on the basis of two structural γ globins, $^{G}\gamma$ and $^{A}\gamma$, differing from each other by one amino acid. $^{G}\gamma$ globin contains a glycine at position 136, whereas $^{A}\gamma$ globin has an alanine at that position. The most stable hemoglobins are tetramers consisting of two α globins and two other globins. The two types of γ globins combine with α globins to form fetal hemoglobin (designated hemoglobin F, HbF, or $\alpha_2\gamma_2$), the major hemoglobin of the fetus; hemoglobin A (HbA, or $\alpha_2\beta_2$) is the major normal adult hemoglobin. Hemoglobin A_2 (HbA_2, or $\alpha_2\delta_2$) is another adult hemoglobin, but is produced at a low level throughout adult life. Two embryonic globin genes, the ζ and ϵ genes, are also identifiable in the fetus, and are part of the embryonic hemoglobins Gower I ($\zeta_2\epsilon_2$), Gower II ($\alpha_2\epsilon_2$), and Portland ($\zeta_2\gamma_2$) (*1*, *4*). The chromosome location of the ζ and ϵ genes is unknown. The molecular events responsible for the switches from embryonic to fetal hemoglobin and from fetal to adult hemoglobin production are also unknown.

0036-8075/80/0201-0486$01.50/0

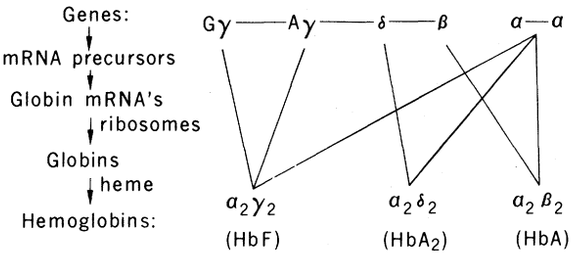

Fig. 1. The major human globin genes and hemoglobins. The four non-α genes are linked in the order shown: $^{G}\gamma$-$^{A}\gamma$-δ-β. HbF is fetal hemoglobin, HbA_2 is hemoglobin A_2, and HbA is normal adult hemoglobin.

The molecular events in the biosynthesis of human globin chains are similar to those for other mammalian globins (*1*). Human globin messenger RNA's (mRNA) have been isolated, and translated in various cell-free systems, including those derived from wheat germ and ascites tumor cells. In these systems, in which the only component of human origin is the mRNA, it is clear that the specificity for the synthesis of human globin is determined by the presence of human globin mRNA, and that translational factors, such as ribosomes, transfer RNA's (tRNA), and other putative regulatory factors at the level of translation, exert relatively minor effects. In other words, the major events regulating the production of human globin are the transcription of the structural human globin genes into nuclear RNA, and subsequent modification of this RNA by RNA processing. In this latter event, the large nuclear RNA is shortened by removal of nucleotides and then methylated at its 5′ end and polyadenylated at its 3′ end to produce mature globin mRNA.

Summary. Studies of the human hemoglobin system have provided new insights into the regulation of expression of a group of linked human genes, the γ-δ-β globin gene complex in man. In particular, the thalassemia syndromes and related disorders of man are inherited anemias that provide mutations for the study of the regulation of globin gene expression. New methods, including restriction enzyme analysis and cloning of cellular DNA, have made it feasible to define more precisely the structure and organization of the globin genes in cellular DNA. Deletions of specific globin gene fragments have already been found in certain of these disorders and have been applied in prenatal diagnosis.

The structural globin gene sequences in RNA and DNA are defined as those required for encoding the amino acid sequence of globin. During the past 2 years, other sequences of DNA within the structural globin genes, but not encoded in the mature globin mRNA, have been found in mouse, rabbit, and human globin genes (*5–8*). These so-called intervening sequences of DNA (or introns) are encoded into nuclear RNA in the mouse, and are presumed to be excised subsequently in the formation of mature globin mRNA (*5*). The process of excision of intervening sequences, and religation of the remaining structural gene sequences to form mature mRNA is called splicing.

The β Thalassemias

Two major types of inherited disorders of hemoglobin exist in man. The first type is represented by qualitative changes in globin, usually resulting from a single nucleotide change leading to a single amino acid change; these disorders are generally referred to as the hemoglobinopathies and are best exemplified by sickle cell anemia, in which both β genes are of the mutant β^S type. The second type is represented by quantitative changes in the amount of normal α or β globin produced, which is best exemplified by a group of disorders known as the thalassemia syndromes. In these diseases, the globins produced have a normal amino acid sequence, but they are reduced in amount. The α and β thalassemias are defined by decreased or absent α or β globin, respectively.

The globin genes are inherited as simple Mendelian traits. For example, heterozygotes for sickle cell trait have one β^S gene and one β^A gene, while sickle cell homozygotes have two β^S genes. Genetic evidence for linkage of the β thalassemia genes to the β globin structural genes is derived from family studies of individuals who are doubly heterozygous for a β thalassemia gene and a β^S gene; only one or another of these alleles is inherited by a single offspring. The heterozygous state of β thalassemia is asymptomatic; the normal β allele compensates for the relative deficiency of β globin biosynthesis due to a β thalassemia gene on the other chromosome. By contrast, in homozygous β thalassemia, there is severe anemia, due to both decreased or absent production of β globin, and a secondary hemolytic process caused by the continued production of α globin chains; these chains are unstable, and precipitate within cells, causing their premature destruction and anemia (*1*). Homozygous β thalassemia, also known as Cooley's anemia, is usually fatal in early adult life because affected patients require multiple blood transfusions. Despite the benefit of these transfusions and the accompanying administration of iron chelating agents, patients usually succumb to the effects of iron toxicity on the heart.

Two general types of β thalassemia, β^+ and β^0 thalassemia, have been characterized. In β^+ thalassemia, there is low-level production of normal β globin from a decreased amount of β mRNA (*9*), even in homozygotes. This indicates that at least one of the β genes is intact in these individuals, and that the defect is due to either defective regulation of transcription of β globin genes or abnormal processing. Varying amounts of β globin biosynthesis in different populations of patients with β^+ thalassemia suggest a heterogeneity of molecular defects responsible for these disorders.

Another relatively common variant of β thalassemia is β^0 thalassemia in which the production of β globin chains in the homozygote is completely absent. Clinically, homozygous β^0 thalassemia is as severe as β^+ thalassemia. Hybridization studies have shown that some β^0 patients have no β globin mRNA, suggesting again that the defect is at the level of defective transcription or abnormal processing or, alternatively, that there may be deletions or base changes in the β globin structural gene (*10–13*). Another group of β^0 thalassemia patients has cytoplasmic β globin mRNA sequences in nearly normal amounts, while others have reduced amounts (*11–13*). At least some of these patients have untranslatable β globin mRNA; in one case, a specific nonsense mutation in the β globin mRNA has been demonstrated (*13*). In a third group of patients with β^0 thalassemia, originating in the Ferrara region of Italy, there is conflicting data as to the defect in β globin mRNA. We have found that β globin mRNA may be incomplete or absent, while others have reported inducible β globin mRNA (*11, 12*).

The β^+ and β^0 thalassemias are to be contrasted with two other related, much rarer disorders: δβ thalassemia and the

hereditary persistence of fetal hemoglobin (HPFH). In these diseases, synthesis of both δ and β globin does not occur at all, and the only hemoglobin synthesized is HbF. In contrast to β^+ and β^0 thalassemias, the anemia in δβ thalassemia is quite mild, and in HPFH is absent.

Fetal Hemoglobin Production in the Thalassemias

The differences in clinical severity between β^+ and β^0 thalassemia on the one hand, and δβ thalassemia and HPFH on the other, are due to differences in the level of compensation by HbF (Table 1). In homozygotes for β^+ and β^0 thalassemia, although the level of HbF production is greater than normal, it is insufficient to compensate for the decreased or absent synthesis of HbA. By contrast, in δβ thalassemia and HPFH, HbF production is increased; in δβ thalassemia compensation is incomplete and a mild anemia is present, whereas in HPFH complete compensation by γ globin usually results in the absence of anemia. These findings indicate differences in the regulation of the γ globin genes in these different conditions, and suggest that, if patients with β^+ and β^0 thalassemia could be made to increase their HbF production as occurs in HPFH, the anemia might be ameliorated.

The absence of HbA_2 in δβ thalassemia and in HPFH first suggested that deletion of nucleotide sequences in the region of the δ globin genes might be responsible for the relative suppression of γ globin synthesis in β^+ and β^0 thalassemia (*14*). Only recently has it been possible to test this hypothesis directly at the molecular level and to ascertain that there are indeed specific deletions of nucleotides in the δ and β globin gene complex in these disorders (*6, 7, 15, 16*); greater deletion occurs in HPFH than in δβ thalassemia, which is consistent with the model that sequences in the δ gene region regulate the relative expression of γ globin genes in these disorders (*6, 7, 12, 16*). The absence of similar deletions in β^+ and β^0 thalassemia further support this hypothesis.

Measuring Globin Genes

The first experiments to support the theory that different extents of deletions in the δβ globin gene complex are responsible for different levels of expression of γ globin genes used liquid hybridization (*12*). In these experiments, specific radioactive probes were prepared to detect and quantify globin genes; these so-called globin complementary DNA (cDNA) probes were synthesized with purified globin mRNA's as templates. In the presence of the enzyme reverse transcriptase and highly radioactively labeled deoxynucleotides, highly specific and sensitive α, β, and γ cDNA probes were obtained, which hybridize (or anneal) specifically with their complement of globin gene sequences in cellular DNA. Under the experimental conditions, the globin cDNA's will not hybridize with nonglobin RNA or DNA sequences, and the α, β, and γ globin cDNA's will only hybridize with homologous nucleotide sequences; for example, α cDNA will not hybridize with β or γ genes, nor will β cDNA hybridize with α or γ genes. However, β globin cDNA will hybridize to δ globin nucleotide sequences because of the close similarity of δ and β globin structural nucleotide sequences. Only 10 of 146 amino acids differ in the δ and β globin chains. Purification of specific globin cDNA's by cloning in bacterial plasmids has provided more specific probes (*17*). This cloned DNA can then be radioactively labeled by the process of nick translation (*18*).

In liquid hybridization experiments, cellular DNA obtained from either spleens or white blood cells of patients with δβ thalassemia and HPFH was shown to have deletions of β-like gene sequences when compared to that of normal patients and those with β^+ and β^0 thalassemia (*12, 15*). Less β cDNA was hybridized to HPFH DNA than to δβ thalassemia DNA, indicating that the deletion of β-like genes in HPFH was greater than that in δβ thalassemia (*12*). These results were consistent with deletion of sequences in both of these disorders with the association of a greater deletion with a more significant amount of γ globin synthesis in HPFH.

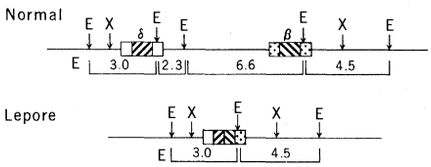

Fig. 2. The linear arrangement of the human δ and β globin structural genes and their flanking and intervening sequences. The 5′ end is at the left. The rectangles represent the globin genes, hatched areas within the genes are the intervening sequences. The single solid horizontal line is the nucleotide sequences between and flanking these genes. The *E*'s and *X*'s with vertical arrows show the sites of cleavage by the enzymes Eco RI and Xba I, respectively. The bracketed areas indicate the DNA fragments generated by Eco RI cleavage; the numbers within the brackets are the number of nucleotides within each fragment in kilobases.

Globin Gene Mapping

In order to more precisely define the molecular defects in the thalassemias and HPFH, it was necessary to analyze the structure and organization of the γ-δ-β gene complex in greater detail and to determine the nature and extent of the deletions and other changes in nucleotide sequences in these disorders. Restriction endonuclease analysis provided the basic technological advance that made this possible (*19, 20*). High-molecular-weight cellular DNA was first isolated from either splenic tissue, white blood cells, or lymphoblast cell lines, and then cleaved by restriction endonuclease at defined nucleotide sequences. For example, the enzyme most frequently used in our studies, Eco RI, only cleaves double-stranded DNA sequences at the hexanucleotide sequences 5′ GAATTC 3′ (G, guanine; A, adenine; T, thymine; C, cytosine). Human DNA contains approximately 3×10^9 nucleotide pairs. Experimentally, the restriction endonuclease Eco RI cleaves human DNA approximately once every 3000 to 5000 nucleotides, and thus generates approximately 10^6 fragments of DNA. The great advantage of using these restriction enzymes is that they cleave each molecule of DNA at precisely the same position, and identical classes of fragments are obtained from each DNA molecule of cellular DNA.

The fragments of cellular DNA resulting from restriction cleavage are separated on the basis of size by agarose gel electrophoresis. The fragments are then transferred in place from the agarose gel to a nitrocellulose filter by a process called blotting, first devised by Southern, and modified by others (*20*). In this procedure, the gel is overlaid with a nitrocellulose filter, and DNA fragments are transferred by blotting induced by dry filter papers on top of the nitrocellulose filter. After transfer, the DNA fragments on the filter are denatured by alkali and hybridized to highly radioactively labeled [^{32}P]globin cDNA under conditions in which only homologous sequences hybridize. The filter is then washed extensively and subjected to radioautography to identify the specific fragments of cellular DNA that have hy-

bridized to the ^{32}P probes. The size and number of these fragments define the cellular DNA fragments containing globin structural gene sequences.

Organization of Normal δ and β Genes

With the use of this technique, the organization of the normal δ and β globin gene complex in human cellular DNA was determined (*6–8*). We knew that β cDNA would hybridize to both δ and β globin structural gene sequences, and that β globin cDNA contained an Eco RI site within the β globin structural sequence (*21*) (Fig. 2). We assumed that this Eco RI site also existed in the δ globin gene DNA, since the amino acid sequence in this region is the same in β and δ globin. On the basis of these assumptions, we could predict that either three or four Eco RI fragments hybridizing to β cDNA would be present, depending on whether there was an Eco RI site between the δ and β genes (Fig. 2). The size of these fragments would be determined by the Eco RI sites flanking the 5′ end of the δ gene and the 3′ end of the β gene. (The 5′ end of the gene is defined as that end of the gene at which transcription of the globin mRNA begins.) In fact, we found four β-like fragments, which hybridized intensely with β globin cDNA (*6, 7*) (Fig. 3).

In order to arrange these four β-like fragments, with respect to each other, we used DNA from a patient homozygous for hemoglobin Lepore. Lepore globin is a fusion gene product, δ-like at its 5′ or NH_2-terminal end, and β-like at its 3′ or COOH-terminal end (*22*). We presumed that the β-like fragments present in both Lepore and normal DNA would define the 5′ δ and the 3′ β gene fragments; the two fragments absent in Lepore DNA and present in normal DNA would represent the 3′ end of the δ gene and the 5′ end of the β gene, and their flanking sequences (Fig. 2). As shown in Fig. 3, two fragments, 3.0 and 4.5 kilobase pairs (kb) in size, are found in both normal and Lepore DNA, and two other fragments, 6.6 and 2.3 kb in size, are found only in normal DNA (*23*). A short cDNA probe that only represented the structural sequence at the 3′ end of the β gene was used to show that the 4.5-kb fragment contained the 3′ end of the Lepore gene, and, therefore, the 3′ end of the β globin gene. Similarly, the 3.0-kb fragment represents the 5′ end of the Lepore gene, and, thus, the 5′ end of the δ gene. The 6.6-kb fragment in normal DNA was shown to contain the 5′ end of the β structural genes (Fig. 2). In other studies, DNA from a patient with a mutant hemoglobin, hemoglobin O Arabia (a β mutant) was used to distinguish the δ and β gene fragments (*8*). In this DNA, a nucleotide change occurs at amino acid codon position 120–121, which destroys the normal Eco RI site in the β gene and gives rise to a single fragment 10 kb in size derived from the 6.6- and 4.5-kb fragments.

Table 1. Hemoglobin defects in thalassemia.

State	Anemia	HbA	HbA_2	HbF
Normal	None	Normal	Normal	Normal
β^+ Thalassemia	Severe	Decreased	Normal	Increased*
β^0 Thalassemia	Severe	Absent	Normal	Increased*
δβ Thalassemia	Mild	Absent	Absent	Increased†
HPFH	None	Absent	Absent	Increased‡

*Slightly increased over normal. †Moderately increased over normal. ‡Markedly increased, that is, completely compensatory.

The physical nucleotide map resulting from these studies showed that the minimum distance between the δ and β structual genes in man is approximately 7 kb. The finding of a single Xba I fragment of approximately 13 kb in size, containing both the δ and β structural genes, proved that these two genes are indeed linked and present on a single piece of DNA (*8*) (Fig. 2). Digestion of DNA with other enzymes alone and in combination led to a more detailed restriction enzyme map of the globin genes.

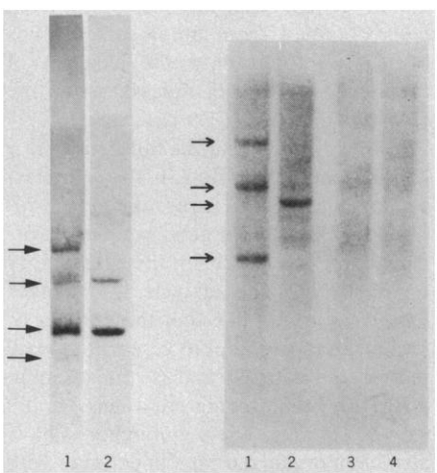

Fig. 3 (left). Autoradiographs of fragments of normal and Lepore DNA's hybridized to ^{32}P-labeled β globin cDNA. Cellular DNA's were digested with Eco RI and the fragments were separated by agarose gel electrophoresis, transferred to nitrocellulose filters, and hybridized to ^{32}P globin cDNA and radioautographed. (Lane 1) Normal DNA generates four fragments containing β-like genes, shown by arrows. (Lane 2) Homozygous Lepore DNA generates only two of the four DNA fragments corresponding to the 5′ δ and 3′ β fragments (3.0 and 4.5 kb in size) shown in Fig. 2. Fig. 4 (right). Autoradiographs of normal, δβ thalassemia, and HPFH DNA's hydridized to ^{32}P-labeled β globin cDNA. Cellular DNA's are digested with Eco RI and fragments are processed as indicated in the legend to Fig. 3. (Lane 1) Normal DNA, only three of the four β-like fragments are seen and are indicated by arrows. The fourth and smallest Eco RI fragment is not seen with this probe. (Lane 2) Homozygous δβ thalassemia DNA; no normal β-like fragments are seen; only a single new β-like fragment shown by the arrow. (Lanes 3 and 4) DNA's from two different patients homozygous for HPFH. No β-like fragments are seen.

Intervening sequences of approximately 1000 base pairs (1 kb in size) were also defined within the δ, β, and Lepore genes by restriction endonuclease analysis (*6–8*). While it was known that β globin cDNA had a single Bam HI restriction site 60 nucleotides 5′ to the Eco RI site in β cDNA, our experiments showed that there was at least 1 kb of DNA between these sites in the cellular δ and β globin genes at a location that is identical to that found in the rabbit and mouse β genes.

The organization of the δ and β structural genes has been confirmed and extended by analysis of these genes after their isolation by cloning of cellular DNA in λ phage (*24, 25*). The location of the large intervening sequences between the Bam HI and Eco RI sites has been more precisely defined, and the size of these sequences was determined as 900 and 950 nucleotides in the δ and β genes, respectively. It has also been shown that there is relatively little sequence homology between these large δ and β intervening sequences. An additional smaller intervening sequence between codons 30 and 31, in both the δ and β genes, has also been discovered in the cloned δ and β genes (*24*).

Deletions in δβ Thalassemia and HPFH

The structure and organization of the δ and β genes was also examined in cell line DNA of patients homozygous for δβ thalassemia and HPFH (*6, 7, 16, 26*). In both cases, we found deletion of all of the normal δ and β globin gene-containing fragments (Fig. 4) (*27*). However, in δβ thalassemia DNA, a new DNA frag-

ment, approximately 4.2 kb in size, is found which hybridizes to β cDNA, consistent with less deletion of β-like genes in $\delta\beta$ thalassemia than HPFH (Fig. 4). This fragment has been shown, with the use of Bam HI, to contain the 5′ end of the δ globin structural gene, and the adjacent 5′ δ flanking sequences (*16, 28*) (Fig. 5). A 15.0-kb Bam HI fragment in normal DNA hybridizes to both β cDNA and γ cDNA, and contains the 5′ end of the δ gene, and the 3′ end of the $^{A}\gamma$ gene (Fig. 5). In homozygous $\delta\beta$ thalassemia DNA, this 15.0-kb fragment persists, whereas other Bam HI fragments, which contain the 5′ and 3′ ends of the β globin structural gene, are missing (*28*). Since $\delta\beta$ thalassemia DNA contains a single 5′ ended β-like sequence, and the 5′ end of the β gene is absent in this DNA, the 15.0-kb fragment must represent the 5′ end of the δ gene and its flanking sequence.

In studies with other restriction enzymes, and a probe representing the region between the δ and γ globin genes, the extent of deletions in $\delta\beta$ thalassemia and HPFH DNA has been further defined (*16*). The deletions in two different patients with HPFH appear to be similar, and extend approximately 3 kb 5′ to the 5′ end of the δ structural gene. These results indicate that the 5′ end of the δ gene and its adjacent sequences are preserved in $\delta\beta$ thalassemia and are absent in HPFH, further supporting the hypothesis that sequences in the region 5′ to the δ gene, or in the δ structural gene itself, are involved in the regulation of expression of the γ globin genes. Alternatively, it is possible that deletion of DNA sequences 3′ to the β gene also occurs in $\delta\beta$ thalassemia and may be important in γ gene expression (*29*).

One reported case appears to argue against a role for γ regulation by sequences 5′ to the δ globin genes (*16, 30*). Restriction endonuclease analysis of the DNA from this individual shows deletion of all δ and β globin gene sequences, as well as of $^{A}\gamma$ sequences on both chromosomes. This patient has the phenotype of $\delta\beta$ thalassemia with anemia, despite deletion of the sequences 5′ to the δ globin genes as in HPFH. An alternative explanation for the decreased γ globin synthesis in this case is that γ globin gene expression also depends on the number of γ globin genes present; when two of the four γ globin genes are deleted, as in this case, maximal γ globin gene expression cannot occur, and the phenotype is that of $^{A}\gamma$ $\delta\beta$ thalassemia. It is of interest in this regard that all homozygotes for HPFH described so far have had both $^{G}\gamma$ and $^{A}\gamma$ synthesis, indicating

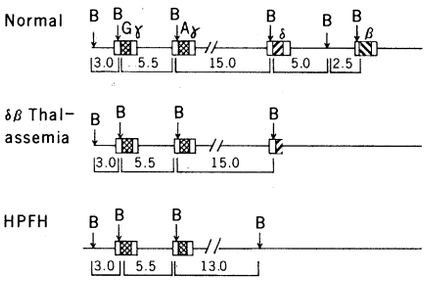

Fig. 5. Arrangement of Bam HI restriction fragments on a physical linear map. The diagrams are similar to those described in the legend to Fig. 1. The Bam HI restriction sites are indicated by B, the vertical arrows; the numbers in brackets below the linear map indicate the sizes of the Bam HI fragments. The normal Bam HI fragments include a 15.0-kb fragment linking the $^{A}\gamma$ and δ globin genes and two smaller fragments, 5.0 and 2.5 kb in size, derived from the δ and β genes. Homozygous $\delta\beta$ thalassemia DNA has the 15.0-kb Bam HI fragment, but lacks the 5.0- and 2.5-kb fragments (*16, 27, 28*). Homozygous HPFH DNA lacks all δ and β gene fragments.

the presence of a full complement of γ globin genes. Analysis of DNA from cells obtained from amniocentesis has been used to confirm the absence of homozygosity for $\delta\beta$ thalassemia in one case (*31*).

The possibility that the differences in γ globin gene expression in the β thalassemias and HPFH are related to differences in γ globin gene content or organization has also been investigated by restriction enzyme analysis. No changes have been found in either the number or organization of γ globin genes in several individuals with β^{+} and β^{0} thalassemia, HPFH, $\delta\beta$ thalassemia, and hemoglobin Lepore (*26, 32*). The γ globin genes have also been organized on a linear map with the use of both restriction endonuclease analysis of cellular DNA and cloning of specific globin genes (*26, 32*). These studies show that the two γ genes are linked to each other and separated by approximately 3.5 kb of nucleotide sequence, and that there is no rearrangement of γ genes in fetal and adult tissues. Each of these genes contains two intervening sequences, a smaller one between codons 30 and 31, and a larger one at the same site as in the δ and β genes between codons 104 and 105. There is strong homology of nucleotide sequences in the intervening sequences of the $^{G}\gamma$ and $^{A}\gamma$ genes. The physical linkage of the two γ genes indicates that these genes are on the same side of the $\delta\beta$ complex (Fig. 5).

The mechanism by which the δ globin gene region or its 5′ flanking sequences can regulate biosynthesis of γ globin is unclear. The possibility exists that, despite evidence for *cis* regulation of γ globin gene expression, RNA transcribed at or near the δ gene locus may play a regulatory role. Despite the lack of significant δ globin production under any circumstances in man, the δ globin structural sequences are homologous to the β structural gene sequences. This homology persists even though there is pronounced divergence of the intervening sequences of the δ and β globin genes. This suggests that perhaps the δ structural gene sequences have a role at the polynucleotide rather than at the polypeptide level. It is possible, for example, that the δ globin mRNA precursors in the nucleus, or transcripts containing these precursors, may affect either γ globin gene transcription or the processing of products including γ globin mRNA precursors. Double-stranded RNA molecules adjacent to structural gene sequences and representing repetitive sequence regions have been discussed as potential sites of regulation (*33*). In addition, it has been shown that a small DNA fragment of approximately 300 nucleotides is highly reiterated in the human genome and has sites of homology in the regions flanking the 5′ end of the δ globin structural gene and 3′ end of the β globin structural gene as well as in the flanking region 5′ to the $^{G}\gamma$ gene and 3′ to the $^{A}\gamma$ gene (*34*). The unique position of these reiterated sequences suggests that they could delimit the origins of RNA transcription of the γ and $\delta\beta$ gene complexes. These sequences are transcribed into nuclear RNA and synthesized in an RNA polymerase III-dependent cell-free in vitro system (*35*). These transcripts preferentially represent the regions 1.5 kb from the 5′ ends of the $^{G}\gamma$ and δ genes. The finding that the repetitive sequence 5′ to the δ gene is deleted in HPFH and not in $\delta\beta$ thalassemia further suggests that this region is involved in the regulation of the γ-δ-β gene complex (*34*).

The intervening sequences also represent potential sites of regulation at the level of transcription or RNA processing. Transcription initiations in genes coding for 5*S* RNA may require specific sequences within the structural genes for their maximal activity (*36*). Of particular interest is the small intervening sequence in the human globin genes, whose location between codons 30 and 31 in the human δ, β, and γ genes is identical to similar intervening sequences in the mouse globin genes. The binding of specific proteins to this region might, for example, be the trigger for initiation of globin mRNA biosynthesis 5′ to this region by permitting the subsequent addition of other proteins including RNA polymerase.

β^+ and β^0 Thalassemia

Detection of the specific nucleotide changes in β^+ and β^0 thalassemia DNA would have considerable import in clinical practice, since these disorders represent a major health problem in certain endemic areas, such as Greece, Italy, and Iran. At present, the prenatal diagnosis of β^+ and β^0 thalassemia requires the use of fetal blood sampling, which is both dangerous to the fetus, in that the mortality rate is 5 to 10 percent, and of limited acceptability to the populations involved. The ability to detect homozygosity for β^+ and β^0 thalassemia in DNA from amniocentesis fluid would permit the precise diagnosis of these disorders with little or no risk. To date, extensive restriction endonuclease analysis of DNA of β^+ and β^0 thalassemia patients has shown no extensive nucleotide deletions in the δ and β gene sequences in the majority of patients studied (*7, 28, 37*). However, in two patients from India, with β^0 thalassemia, a 0.8-kb deletion of the 3′ end of the β gene has been found (*37, 38*). It is of interest that in these cases only one of the two β^0 thalassemia genes has the deletion defect; the other gene appears to be largely structurally intact; thus, even in these populations, there is heterogeneity of the genetic defect in these syndromes.

From these data, it appears that the DNA defect in most patients with β^+ and β^0 thalassemias is small, and may involve only a single nucleotide base change. In one case of β^0 thalassemia, a single nucleotide change in codon 17, causing an amber mutation, has been demonstrated (*13*). Recent data suggest that at least some β^+ thalassemias are due to DNA defects in the intervening sequences leading to abnormal RNA processing (*39*). Further analysis of RNA synthesis in these patients and the ability to clone and sequence the β globin genes in β^+ and β^0 thalassemia DNA should further define the gene defect in these disorders (*25*). The rapid improvement in this so-called shotgun cloning, with the use of unfractionated human DNA and λ phage as the vector, permits analyses of this type. Clones are already available from patients with β^+ thalassemia, which contain the entire δ and β globin gene complex and the sequences flanking these genes (*40*) while other clones contain parts of these genes (*41*). The nucleotide sequence of these clones can be compared directly to those in clones of δ and β genes from nonthalassemia patients. It may be necessary to compare several different clones of β^+ and β^0 thalassemia DNA with those from different normal DNA's to distinguish nucleotide changes that are important for β globin gene function, from other polymorphisms in human DNA that do not influence globin gene function.

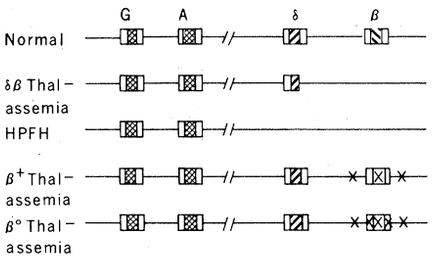

Fig. 6. Summary of defects in the β thalassemias and related disorders. The normal genes are shown at the top and as in Figs. 1 and 5. The deletions in $\delta\beta$ thalassemia and HPFH are shown. The postulated defects in β^+ and β^0 thalassemia DNA are indicated by crosses.

The heterogeneity of patients with β^+ and β^0 thalassemia, both clinically and biochemically, suggests that many different defects may be found at the molecular level. Thus, these disorders provide examples of several mutations in DNA, which can result in a common functional abnormality in the biosynthesis of a specific protein (Fig. 6). In this sense, the β thalassemias represent a model of a group of diseases that may reveal the nucleotide sequences involved in the regulation of a specific protein.

The α Thalassemias

The α thalassemias are characterized by a decrease in or absence of α globin biosynthesis. Since in most populations there appear to be two α globin genes per chromosome, there are four putative states of α thalassemia characterized by the loss of one to four α globin gene loci (*42, 43*). Studies with α globin cDNA indicate that the majority of the α thalassemias are due to deletions of α gene loci. When one α globin gene is deleted, the so-called silent carrier state for α thalassemia exists manifested only by decreased α globin biosynthesis. When two of four α genes are deleted, the α thalassemia trait is present, associated with morphologic abnormalities of the red cells, but no significant anemia. Deletion of three of four α genes results in hemoglobin H disease; hemoglobin H is a tetramer composed of β chains, which accumulates because of the lack of adequate amounts of α globin. Hemoglobin H is an unstable hemoglobin which precipitates in cells and causes a mild hemolytic anemia. The loss of all four α globin genes is associated with death of the fetus in utero in a disease called hydrops fetalis-α thalassemia, a syndrome characterized by congestive heart failure in the fetus due to severe anemia.

Restriction endonuclease mapping has confirmed the loss of α globin loci associated with the various forms of the α thalassemia syndromes (*44*). It appears that in Oriental populations both α globin genes are commonly deleted from a single chromosome; thus, the presence of hydrops fetalis is explained; in contrast, it is unusual for two α globin genes to be lost from a single chromosome in Mediterranean or black populations, and the incidence of hydrops fetalis in these populations is negligible. It is common, however, for one α globin gene to be deleted on each of two chromosomes in black individuals, making the incidence of silent carrier trait and α thalassemia trait significant in this group.

The prenatal diagnosis of α thalassemia can be made by analysis of DNA from amniocentesis fluid cells (*43*). Its major clinical usefulness is that it permits a mother, with a fetus with hydrops fetalis, to choose whether to carry the fetus through the full 9 months of pregnancy. The definition of other forms of α thalassemia prenatally is unnecessary, since all of the other states are compatible with normal life.

Molecular Genetic Polymorphisms

While most studies of DNA in disorders of globin biosynthesis have focused on determining the defect in nucleotide sequence of the DNA within or surrounding the globin genes, it has been shown that changes in DNA at a distance from the structural genes can be used to detect the presence of these genes. Kan and Dozy (*45*) discovered a strong association of the sickle cell (β^S) gene with a polymorphism in the Hpa I restriction enzyme site surrounding the β^S gene. In normal DNA, from all Caucasian patients studied, and from a majority of black patients with no sickle cell genes, a 7.6-kb restriction fragment is generated by Hpa I; this fragment contains the entire β globin gene.

By contrast, it was originally reported that the β^S gene is associated with a 13.0-kb β globin fragment about 87 percent of the time. This is due to either a change in the Hpa I site 3′ to the β structural gene, or to an insertion of sequences in this region. The association of the 13.0-kb Hpa I fragment with the β^S gene has been used for the prenatal diagnosis of sickle cell anemia by analyzing DNA

from amniocentesis fluid, and demonstrating the presence of only the 13.0-kb fragment (*45*). Studies in our laboratory have largely confirmed these observations except that the incidence of the association of the β^S gene with the 13.0-kb fragment has been found to be closer to 60 percent (*46*).

It has also been found that the association of the 13.0-kb fragment is by no means specific for the β^S gene (*46*, *47*). In individuals with hemoglobin C ($\alpha_2\beta^C{}_2$), the association of the β^C gene with 13.0-kb fragment is approximately the same as with the β^S gene, 50 to 60 percent. Individuals with only β^A genes have also been identified with the 13.0-kb fragment, with a frequency of between 5 and 10 percent. Taken together, these data indicate that the 13.0-kb fragment is not specifically linked to the β^S gene, but is more likely the result of a polymorphism in a subpopulation of blacks who subsequently developed the β^S and β^C genes by mutation. Whether other nucleotide changes exist in the region surrounding the Hpa I site in sickle cell anemia, which might be additionally useful for diagnosis of this disorder, remains to be determined. An association between a specific Bam HI fragment and the β^0 thalassemia gene has been used in the Sardinian population for prenatal diagnosis (*48*). The study of genetic polymorphisms may also be useful in the diagnosis of other inherited disorders if specific linkages between the genes being screened and the polymorphisms detected can be established.

Future Approaches

The specific methodology used in the analysis of the human globin gene system can be applied to other genes in man as well. The primary requirement for this type of analysis is a probe for recognizing specific gene sequences. With the availability of clones containing complete structural genes and their flanking sequences, it should be possible to begin to reconstruct systems capable of gene transcription. In these systems, purified genes can be used to bind and isolate specific nuclear proteins. The recent description of in vivo systems for gene transfer and gene transcription makes the availability of cloned fragments of DNA of even greater interest (*49*). At long range, it may be possible to use isolated human globin genes to transform human bone marrow cells, and these genes may integrate and express newly acquired normal globin gene function. This would be the ideal therapy for patients with sickle cell anemia and homozygous β thalassemia.

Conclusions

Restriction endonuclease analysis and gene cloning have made it possible to determine the physical map of the human globin genes. The linkage of the γ, δ, and β globin genes, the distance between these genes, and the presence of intervening sequences within each have been established. In certain disorders, such as $\delta\beta$ thalassemia and HPFH, the extent of deletion of δ and β globin gene fragments has been mapped and shown to be directly related to the relative expression of γ globin genes. From these studies, it appears that a specific region in the 5′ end of the δ structural gene or its flanking 5′ sequences (or both) may be important in the regulation of expression of γ globin genes.

By contrast, in the β^+ and β^0 thalassemias, no deletion of the γ-δ-β globin gene complex has yet been found, and the relative amount of γ globin gene compensation is low, resulting in severe anemia. These diseases are most likely due to either decreased transcription of β globin mRNA or abnormal processing of β globin mRNA precursors. It is now possible, by means of recombinant DNA cloning technology, to obtain clones of β^+ and β^0 thalassemia genes for restriction endonuclease analysis, and nucleotide sequencing.

In the α thalassemias, deletions of α globin genes have been demonstrated both by liquid hybridization studies and by restriction endonuclease analysis of DNA. The prenatal diagnosis of α thalassemia can be made by analysis of amniocentesis fluid DNA. In sickle cell anemia, a distinctive polymorphism in an Hpa I cleavage site has been used to define the presence of the β^S gene in 60 to 80 percent of patients with this gene. In families in which this polymorphism is present, it is possible to make the prenatal diagnosis of β sickle cell anemia by amniocentesis.

References and Notes

1. D. J. Weatherall and J. B. Clegg, *The Thalassemia Syndromes* (Blackwell, Oxford, ed. 2, 1972); H. F. Bunn, B. G. Forget, H. M. Ranney, in *Hemoglobinopathies*, L. H. Smith, Ed. (Saunders, Philadelphia, 1977), vol. 12, p. 28; A. Bank, *Blood* **51**, 369 (1978).
2. A Deisseroth, A. Nienhuis, J. Lawrence, R. Giles, P. Turner, F. H. Ruddle, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1456 (1978).
3. A. Deisseroth *et al.*, *Cell* **12**, 205 (1977).
4. R. E. Gale, J. B. Clegg, E. R. Huehns, *Nature (London)* **280**, 162 (1979).
5. A. J. Jeffries and R. A. Flavell, *Cell* **12**, 1097 (1977); S. M. Tilghman *et al.*, *Proc. Natl. Acad. Sci, U.S.A.* **75**, 725 (1978); S. M. Tilghman, P. J. Curtis, D. C. Tiemeier, P. Leder, C. Weissman, *ibid.*, p. 1309.
6. J. G. Mears *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1222 (1978).
7. J. G. Mears, F. Ramirez, D. Leibowitz, A. Bank, *Cell* **15**, 15 (1978).
8. R. A. Flavell, J. M. Kooter, E. DeBoer, P. F. R. Little, R. Williamson, *ibid.*, p. 25.
9. A. Bank and P. A. Marks, *Nature (London)* **212**, 1198 (1966); *J. Clin. Invest.* **45**, 330 (1966); J. B. Clegg, M. A. Naughton, D. J. Weatherall, *J. Mol. Biol.* **19**, 91 (1966); A. Nienhuis and W. F. Anderson, *J. Clin. Invest.* **50**, 2458 (1971); E. J. Benz and B. G. Forget, *ibid.*, p. 2755; L. W. Dow *et al.*, *Nature (London)* **243**, 114 (1973); D. L. Kacian *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1886 (1973); D. Housman, B. G. Forget, A. Skoultchi, E. J. Benz, *ibid.* p. 1809; F. Ramirez *et al.*, *ibid.* **72**, 1550 (1975); U. Nudel, F. Ramirez, P. A. Marks, A. Bank, *J. Biol. Chem.* **52**, 2182 (1977); N. S. Young, E. J. Benz, J. A. Kantor, P. Kretschmer, A. W. Nienhuis, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5884 (1978).
10. B. G. Forget, E. J. Benz, A. Skoultchi, C. Baglioni, D. Housman, *Nature (London)* **247**, 379 (1974); Y. W. Kan, J. P. Holland, A. M. Dozy, H. E. Varmus, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 5140 (1975); F. Ramirez, D. Starkman, A. Bank, H. Kerem, G. Cividalli, E. A. Rachmilewitz, *Blood* **52**, 7535 (1978).
11. S. Ottolenghi, P. Comi, B. Giglioni, R. Williamson, C. Vullo, L. Del Senno, F. Conconi, *Nature* **266**, 231 (1977).
12. F. Ramirez *et al.*, *ibid.* **263**, 471 (1976).
13. J. C. Chang and Y. W. Kan, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2886 (1979).
14. T. H. J. Huisman *et al.*, *Ann. N.Y. Acad. Sci.* **232**, 107 (1974).
15. Y. W. Kan, J. P. Holland, A. M. Dozy, H. E. Varmus, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 5140 (1975); B. G. Forget *et al.*, *Cell* **7**, 232 (1976); S. Ottolenghi *et al.*, *ibid.* **9**, 72 (1976).
16. E. F. Fritsch, R. M. Lawn, T. Maniatis, *Nature (London)* **279**, 598 (1979); S. Ottolenghi *et al.*, *ibid.* **278**, 654 (1979).
17. T. Maniatis, G. K. Sim, A. Efstradiadis, F. C. Kafatos, *Cell* **8**, 163 (1975); F. R. Blattner *et al.*, *Science* **196**, 161 (1977); J. T. Wilson *et al.*, *Nucleic Acids Res.* **5**, 563 (1978); P. F. R. Little *et al.*, *Nature (London)* **272**, 640 (1978).
18. T. Maniatis, A. Jeffrey, D. G. Kleid, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1184 (1975).
19. D. Nathans and H. O. Smith, *Annu. Rev. Biochem.* **44**, 273 (1975); R. J. Roberts, *CRC Crit. Rev. Biochem.* **4**, 123 (1976).
20. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1876); M. Botchan, W. Topp, J. Sambrook, *Cell* **9**, 269 (1976); A. J. Jeffries and R. A. Flavell, *ibid.* **12**, 429 (1977).
21. C. A. Marotta, J. T. Wilson, B. G. Forget, S. M. Weissman, *J. Biol. Chem.* **252**, 4050 (1977).
22. C. Baglioni, *Proc. Natl. Acad. Sci. U.S.A.* **48**, 1880 (1962).
23. The size of DNA fragments has varied between 0.5 and 1.0 kb in different reports depending on the amount of DNA and the particular electrophoresis conditions used. The largest differences have been between cellular DNA fragments and cloned DNA's; for example, our 4.5-kb fragment has been determined as 3.2 kb, our 3.0-kb fragment as 2.3 kb, and the 6.6-kb fragment as 5.2 kb by others (*24*). The organization of fragments with relation to each other has been quite consistent.
24. R. M. Lawn, E. F. Fritsch, R. C. Parker, G. Blake, T. Maniatis, *Cell* **15**, 1157 (1978).
25. T. Maniatis *et al.*, *ibid.*, p. 687.
26. D. Tuan, P. A. Biron, J. K. deRiel, H. Lazarus, B. G. Forget, *Nucleic Acids Res.* **6**, 2519 (1979); O. Smithies *et al.*, *Science* **202**, 1284 (1978).
27. The 4.5-kb fragment initially described as a residual fragment that contains a β-like gene in $\delta\beta$ thalassemia and in HPFH (*6*) has not been observed when cloned purified β cDNA is used. This fragment probably is an embryonic gene fragment hybridizing with the cDNA used in these experiments.
28. A Bank, J. G. Mears, F. Ramirez, A. L. Burns, J. Feldenzer, in *Eucaryotic Gene Expression*, R. Axel, T. Maniatis, F. C. Fox, Eds. (New York Academy of Sciences, New York, 1979).
29. R. Bernards, J. M. Kooter, R. A. Flavell, *Gene* **6**, 265 (1979).
30. S. H. Orkin, B. P. Alter, C. Altay, *J. Clin. Invest.* **64**, 866 (1979).
31. ———, J. Mahoney, H. Lazarus, J. C. Hobbins, D. C. Nathan, *N. Engl. J. Med.* **299**, 166 (1978).
32. P. F. R. Little, R. A. Flavell, J. M. Kooter, G. Annison, R. Williamson, *Nature (London)* **278**, 227 (1979); F. Ramirez, A. L. Burns, S. Spence,

J. G. Mears, A. Bank, *Nucleic Acids Res.* **7**, 114 (1979); O. Smithies, personal communication.
33. E. H. Davidson and R. J. Britten, *Science* **204**, 1052 (1979).
34. W. Jelinek, R. Evans, M. Wilson, M. Salditt-Georgieff, J. E. Darnell, *Biochemistry* **17**, 2776 (1978); L. Lacy and T. Maniatis, personal communication.
35. C. Duncan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5095 (1979).
36. S. Sakonju, D. Bogenhagen, D. D. Brown, personal communication.
37. R. A. Flavell *et al.*, *Nucleic Acids Res.* **6**, 2749 (1979).
38. S. H. Orkin, J. M. Old, D. J. Weatherall, D. G. Nathan, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2400 (1979).
39. N. S. Young, E. J. Benz, J. A. Kantor, P. Kretchmer, A. W. Nienhuis, *ibid.* **75**, 5884 (1978); L. E. Mequat and J. Ross, personal communication.
40. A. L. Burns, S. Spence, F. Ramirez, J. G. Mears, A. Bank, *Blood* **54**, 51a (1979).
41. R. A. Spritz *et al.*, *ibid.*, p. 60a.
42. J. M. Taylor *et al.*, *Nature (London)* **251**, 392 (1974); S. Ottolenghi *et al.*, *ibid.*, p. 389.
43. Y. W. Kan, M. S. Golbus, A. M. Dozy, *N. Engl. J. Med.* **295**, 1165 (1977).
44. S. H. Orkin, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5950 (1978); _____, J. Old, H. Lazarus, C. Altay, A. Gurgey, D. J. Weatherall, D. G. Nathan, *Cell* **17**, 33 (1979); S. H. Embury, R. V. Lebo, A. M. Dozy, Y. W. Kan. *J. Clin. Invest.*, in press.
45. Y. W. Kan and A. M. Dozy, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5631 (1978); *Lancet* **1978-II**, 910 (1978).
46. J. Feldenzer, J. G. Mears, A. L. Burns, C. Natta, A. Bank, *J. Clin. Invest.* **64**, 751 (1979).
47. J. C. Chang, G. F. Temple, R. F. Trecartin, Y. W. Kan, *Clin. Res.* **27**, 457 (Abstr.) (1979).
48. Y. W. Kan, K. V. Lee, M. Furbetta, A. Angius, A. Tao, *Blood* **54**, 55a (1979).
49. M. Wigler *et al.*, *Cell* **16**, 777 (1979); M. Wigler *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1373 (1979).
50. Supported by NIH grants GM-23143-04, CA-13696, and AM-25274-13, NSF grant 24120, a grant from the National Foundation-March of Dimes, and a grant from the Cooley's Anemia Foundation.

R & D Employment in the U.S.S.R.

Louvan E. Nolting and Murray Feshbach

In the face of labor shortages, growing inaccessibility of raw materials, rising capital needs and defense expenditures, and increasing economic complexity, the Soviet Union is relying heavily on scientific and technical progress to sustain its economic growth. Employment in research and development (R & D) is a prime indicator of a nation's reliance on science and innovation. This article examines the number and distribution of persons engaged in R & D in the U.S.S.R. and makes quantitative comparisons between Soviet and U.S. R & D employment.

The two main Soviet statistical series which relate to R & D personnel are the series on employment in the science and science services branch of the economy and the series on scientific workers. The first series covers the branch definition of employment in specialized scientific organizations engaged in R & D and science services. The second covers the functional definition by including all persons with advanced scientific degrees and titles, teachers in higher educational institutions, research managers, and highly educated scientists and engineers whose primary activity is research or general R & D under science plans approved by higher authority. Employment according to the branch definition is more than three times greater than that according to the functional definition. The first series serves in this article as the basis for estimating the number of personnel engaged in conducting and supporting various kinds of R & D; the second series is used to calculate personnel directly conducting R & D. Neither series is wholly comparable with Western data on R & D personnel, because they include some non-R & D employment and exclude some participants in R & D (Table 1).

Employment in Organizations

Conducting and Supporting R & D

The science and science services branch of the Soviet national economy embraces most of what are called scientific organizations (specialized organizations not incorporated into other enterprises or social institutions), which conduct or provide services for scientific research, development, and innovation, presumably in both the civilian and military sectors of the economy.

Scientific organizations are divided into two categories: (i) scientific institutions and (ii) a wide variety of organizations that do not carry out scientific research but are involved in development (in Soviet terminology, test-design work), in production and testing of prototypes and development of new processes, and in the provision of data and services required by scientific institutions. The first category includes only those institutions that (i) systematically conduct scientific research work in a scientific field, including the social sciences and humanities, (ii) work under science plans officially approved by higher authority, (iii) are funded under planned financing for science, and (iv) are staffed by scientific workers (*1*). Scientific institutions include scientific research institutes and laboratories; design organizations, industrial testing stations, and agricultural experimental stations engaged in research; and a variety of organizations—higher educational institutions, museums, observatories, zoos, natural parks, computer centers, libraries, and state archives—that conduct research in addition to their primary functions. Scientific organizations not classified as scientific institutions include nonresearch design organizations, experimental plants that do not produce industrial products for sale, maritime resources prospecting organizations, industrial testing laboratories, computer and information centers not engaged in research, hydrometeorological organizations, and geological survey and prospecting organizations.

Table 2 gives the total number of Soviet scientific institutions and their distribution by type. The number of scientific institutions has increased from 3447 in 1950 to 5327 in 1975, or by more than 50 percent. The increase was primarily a reflection of a twofold increase in the number of scientific research institutes. The number of independent design bureaus

Louvan E. Nolting and Murray Feshbach are Social Science Analyst and Chief, U.S.S.R. Population Employment and R & D Branch, respectively, in the Foreign Demographic Analysis Division, U.S. Bureau of the Census, Washington, D.C. 20233. Dr. Feshbach is presently on leave as a fellow of the Woodrow Wilson International Center for Scholars. This article is a condensation of a report published in 1979 by the Joint Economic Committee of the U.S. Congress and of a larger forthcoming report by the Foreign Demographic Analysis Division, Bureau of the Census, U.S. Department of Commerce. Many of the statistics and definitions were obtained as part of the U.S./U.S.S.R. Agreement on Science and Technology through participation in the Science Policy subgroup on Training and Utilization of Scientific-Technical Manpower.

0036-8075/80/0201-0493$02.00/0