Volume 6 Number 7 1979 Nucleic Acids Research

Restriction endonuclease mapping of the human γ globin gene loci

Dorothy Tuan, P.Andrew Biro, Jon K.deRiel, Herbert Lazarus* and Bernard G.Forget

Departments of Internal Medicine and Human Genetics, Yale University School of Medicine, New Haven, CT 06510 and *The Sidney Farber Cancer Institute, Boston, MA 02115, USA

Received 4 April 1979

ABSTRACT

The restriction endonuclease sites in and around the human γ globin gene loci have been mapped using the gel blotting technique of Southern, in both normal DNA and DNA from an individual with hereditary persistence of fetal hemoglobin (HPFH). In normal DNA, the γ genes are linked to the δ (and β) globin genes, and the orientation of these genes with respect to transcription is (5') $G_\gamma \rightarrow A_\gamma \rightarrow \delta \rightarrow \beta$ (3'). The distance between the G_γ and A_γ genes is 3.5 kb and that between the A_γ and δ genes is 16 kb. In both normal DNA and HPFH DNA, the γ genes are interrupted by an intervening sequence, approximately 1 kb in length that is situated between codon positions 99 and 121 of the coding sequence. In different DNA samples, there is polymorphism for the presence or absence of a Hind III site in the intervening sequence of either γ globin gene. In HPFH DNA, a deletion of at least 16 kb of DNA has been detected. This deletion starts at a point approximately 12.5 kb from the 3'-end of A_γ gene and extends through the δ and β globin genes to a point at least 3 kb beyond the 3'-end of β globin gene.

INTRODUCTION

It has been estimated from genetic analysis and from molecular hybridization analysis that the human genome contains no fewer than four non-α globin genes: one δ gene, one β gene, and two (or more) γ globin genes coding for the two different γ globin chains that differ by the presence of either alanine ($^A\gamma$) or glycine ($^G\gamma$) at amino acid position no. 136 (for reviews see Ref. 1 & 2). From genetic analysis of variants resulting from non-homologous crossing over between these genes, it has been determined that these genes are linked on the same chromosome, and that the gene arrangement in the direction of transcription is $^A\gamma \rightarrow \delta \rightarrow \beta$; from genetic analysis alone, the $^G\gamma$ gene could be situated to either side of this cluster. In spite of their linkage, the expression of these genes is not synchronous: the γ genes are maximally expressed during fetal development, but in adult life γ globin chain synthesis is only minimal. Although the δ and β genes are expressed synchronously in adult life, their levels of expression are quite different: the δ globin gene

product accumulates to a level of only 2% to 3% of that of the β globin gene (1,2).

In an effort to understand the molecular mechanisms that may be associated with the differential activation and expression of the human non-α globin genes, we have mapped the restriction endonuclease cleavage sites in and around the γ gene loci in both normal DNA and DNA from an individual with hereditary persistence of fetal hemoglobin (HPFH). HPFH DNA is of interest, because in individuals with HPFH, γ globin gene expression persists at a high level into adult life, and adequately compensates for the absence of δ and β globin chain synthesis that is associated with deletion of the δ and β globin genes (3,4). Since the elements controlling eukaryotic gene expression are thought to be encoded in DNA sequences contiguous to the structural genes (5) a comparison of the sequences flanking the γ globin genes in HPFH and normal DNA could point to a region of DNA, the deletion of which might be responsible for the persistent expression of γ globin genes in HPFH.

In these studies, we have used restriction endonuclease enzymes such as Bam HI, Eco RI and Pst I that have known cleavage sites within the coding sequences of one or both of the γ globin genes (6,7) as well as enzymes such as Bgl II, Xba I, and Hind III, that do not cut within the γ coding sequences. The intragenic Bam HI site at codon positions 99-100 and the Eco RI site at codon positions 121-122 have been used as reference points in double digestions of DNA to map the cleavage sites for the enzymes Bgl II, Hind III and Xba I around the γ globin genes. We have used γ globin cDNA plasmid clones JW151 (8), pHγG1 (9) and pRP10 (7) as probes for the γ globin genes, and subclones from the δ-β genomic DNA clone HβG1 (10) as probes for δ and β globin gene sequences. A subclone of DNA fragment H (10), derived by Eco RI cleavage of HβG1 and located approximately 5 kb to the 5'-side of the δ globin gene, was used as a probe for the inter γ-δ globin gene region.

Employing the method of Southern (11) with DNA singly and/or doubly digested with various restriction enzymes, we have confirmed the orientation of the human non-α globin genes to be (5') $^{G}\gamma\rightarrow{}^{A}\gamma\rightarrow\delta\rightarrow\beta$ (3'), and have determined the inter-gene distances between the $^{G}\gamma$, $^{A}\gamma$ and δ globin genes. In addition, the existence as well as the length of large intervening sequences interrupting the γ globin genes have been established, as well as the minimum extent and the approximate 5'-end point of the DNA deletion in HPFH DNA.

MATERIALS AND METHODS

(a) Preparation and Digestion of Cellular DNA with Restriction Endonucleases

Normal DNA was prepared from the placenta of a black child obtained at the time of childbirth. HPFH DNA was prepared from the lymphoblastoid cell line established from peripheral blood cells of an individual with HPFH (4). Other samples of non-HPFH DNA were obtained from the placenta of an Asian infant with homozygous α-thalassemia, the peripheral blood leukocytes of a black patient with sickle cell anemia, the lymphoblastoid cell line of a Mediterranean individual with homozygous δβ thalassemia (12) and the spleen of a Caucasian individual with immunohemolytic anemia. High molecular weight DNA was prepared from the above tissues by the method of Blin and Stafford (13).

Restriction endonuclease enzymes were obtained from New England Biolabs Inc. DNA was digested overnight using 5 to 10 μg of DNA/unit of enzyme in the buffer and at the temperature recommended by the enzyme manufacturer. When a second restriction endonuclease digestion was required, the singly digested DNA was extracted twice with phenol, precipitated from ethanol then redissolved in the appropriate buffer for digestion with the second restriction enzyme.

(b) Agarose Gel Electrophoresis

Gels of 0.7% to 1.2% agarose, 20 cm x 20 cm x 0.6 cm, were cast in 80 mM Tris, pH 7.8, 40 mM sodium acetate and 4 mM EDTA (electrophoresis buffer), in a horizontal gel apparatus (14). After digestion with restriction enzymes, DNA samples were extracted twice with phenol,precipitated from ethanol and redissolved in 200 μl of electrophoresis buffer. The DNA samples were then mixed with 100 μl of 1% agarose and loaded into sample wells (3 cm x 0.2 cm x 0.5 cm); gels were run at room temperature for 18 hours at 50V. In order to avoid smearing of the DNA bands, the amount of DNA loaded was kept at less than 1.5 $\mu g/mm^2$ of cross-sectional area of the gel. Phage λ cI857 DNA cut with Hind III or Eco RI, and labeled with [γ-^{32}P] ATP (New England Nuclear, 2000 to 3000 Ci/m mol) and polynucleotide kinase (15) served as the source of size markers (16). Since the mobility of DNA in agarose gels varies with the amount (or concentration) of the DNA that is loaded on the gel, ^{32}P-labeled size markers (0.2-0.4 μg) were co-electrophoresed with an amount of digested human cellular DNA equivalent to that under study. We have found that reference to size markers from Hind III or Eco RI cut phage λ cI857 DNA visualized by florescent staining with ethidium bromide (which requires 2 to 4 μg of λ DNA) leads to an underestimate of the sizes of human globin gene fragments in gel blotting experiments, because of the higher DNA concentration and therefore slower migration of each size marker.

(c) Transfer of DNA to Nitrocellulose Filters

After electrophoresis, the DNA in the agarose slab gels was denatured in 1 M NaCl+0.5 M NaOH for 45 minutes; following neutralization of the solution with 1.5 M NaCl+0.5 M Tris, pH 7, the DNA was transferred by blotting of the gel overnight onto nitrocellulose filters (Schleicher and Schuell, Inc., 0.45 μm pore size) according to the procedure of Southern (11) in the presence of 20x SSC. Following the transfer, the nitrocellulose filters were rinsed in 2x SSC, air dried and baked for 3 to 5 hours at 80°C in a vacuum oven.

(d) Filter Hybridization

Baked filters were soaked in 0.02% Ficoll, 0.5% polyvinylpyrrolidone, 0.02% bovine serum albumin, 0.1% SDS, and 2x SSC for 3 to 5 hours at 63° to 65°C. The filters, cut into 0.7 cm x 20 cm strips, were then incubated for 2 days at 63° to 65°C in sealed plastic bags with 0.5 to 1.0 ml of hybridization solution containing, in addition to the above ingredients, 10 μg/ml of poly(A), 50 μg/ml of denatured salmon sperm DNA, and 0.1 to 0.2 μg/ml of heat denatured globin gene probe previously labeled with ^{32}P by nick-translation (17) to a specific activity of 10^7 to 10^8 cpm/μg. After hybridization, the strips were washed five times over a period of 3 hours at 65°C with hybridization solution lacking ^{32}P-labeled globin probe, poly(A), and denatured salmon sperm DNA. The strips were then rinsed twice for 15 minutes with a solution of 0.1x SSC and 0.1% SDS at 63° to 65°C (18) and rinsed a final time with 2x SSC at room temperature. The strips were then air dried and autoradiography accomplished by exposure for one to three weeks at -20°C to Kodak XR-1 film with Dupont Cronex lightning-plus intensifying screens.

(e) Hybridization Probes

Cloned γ globin cDNA plasmids JW151 (8) or pHγG1 (9) (the latter kindly provided by Dr. S. Malcolm) were used as γ globin gene probes. Probes for the δ and β globin genes were obtained from the genomic δ-β gene clone HβG1 (10); individual DNA fragments containing the δ and β globin genes with adjacent flanking sequences, were obtained by digestion of HβG1 with Pst I, were subcloned in pBR322 by Dr. T. Maniatis and co-workers, and were kindly provided to us by Dr. Maniatis. To determine the γ to δ intergene distance as well as to map DNA sequences in the region between the γ and δ globin genes, we used as probe a pMB9 plasmid subclone (also kindly provided by Dr. T. Maniatis) that contains the Eco RI fragment H of HβG1 (10) consisting of 0.5 kb of DNA located 4 to 5 kb to the 5'-side of the δ globin gene.

cDNA plasmid probes containing sequences complementary to the 5'- and 3'-ends of γ mRNA were prepared by Eco RI + Hind III double digestion of plasmid pHγG1 (19). The DNA fragments were separated by agarose gel electrophoresis

and the DNA extracted by grinding the appropriate gel slices in 0.1x SSC; the eluted DNA (in the supernatant solution after centrifugation) was recovered by precipitation with ethanol. We have also used as a 3'-γ probe, a γ cDNA clone pRP10 (7), kindly provided by Drs. R. Poon and Y.W. Kan, that contains sequences complementary to $^{A}\gamma$ mRNA from codon position no. 99 to the 3'-end of the mRNA. A probe containing sequences complementary to the 5'-end of the δ gene was prepared from the pBR322 δ gene subclone (Hδ1) by double digestion of the plasmid DNA with Pst I + Bam HI.

The various probes were labeled with ^{32}P to a specific activity of 10^7 to 10^8 cpm/μg by nick-translation (17) at room temperature for 2 to 4 hours using [α-^{32}P] deoxyribonucleotide triphosphates (dCTP and TTP, Amersham Corp., 300 Ci/m mol) and DNA polymerase I (Boehringer Mannheim Biochemicals). Before use, the nick-translated probes were passed through a column of Sephadex G-50 to separate the DNA probe from the unincorporated triphosphates.

RESULTS AND DISCUSSION

(a) Bam HI Cleavage Sites in and around the γ Globin Genes - Linkage of the two γ Globin Genes to one another and to the δ Globin Gene

From direct sequence analysis of human γ globin cDNA plasmids, it has been established that there is a Bam HI cleavage site at codon positions 98 to 100 of the $^{G}\gamma$ globin gene (6); a similar site is in all likelihood present at the same position in the $^{A}\gamma$ gene although it has not yet been directly demonstrated by sequence analysis. When normal DNA is digested with Bam HI, three DNA fragments are obtained that hybridize to γ globin cDNA and have sizes of 18 kb, 5.1 kb and 2.6 kb, respectively. The 18 kb fragment hybridizes only to the γ cDNA probe complementary to the 3'-end of γ mRNA; the 2.6 kb fragment hybridizes only to 5'-γ cDNA probe, whereas the 5.1 kb fragment hybridizes to both 5'- and 3'-γ cDNA probes(Figure 1). If there are only two γ globin genes per haploid genome (one $^{A}\gamma$ gene and one $^{G}\gamma$ gene) as suggested by DNA-cDNA hybridization assays (20,21), as well as by restriction endonuclease mapping (19), then the 5.1 kb band is the DNA fragment that spans the two γ globin genes and encompasses the DNA between the two intragenic Bam HI sites of the γ genes. The orientation of the three Bam HI fragments with respect to transcription is therefore (5') 2.6 kb → 5.1 kb → 18 kb (3'). It is not possible, from Bam HI digestion alone, to identify which of the γ genes ($^{A}\gamma$ or $^{G}\gamma$) is situated transcriptionally to the 5'-side of the γ globin gene cluster. For convenience, we shall refer to the γ gene situated on the 5'-side as the

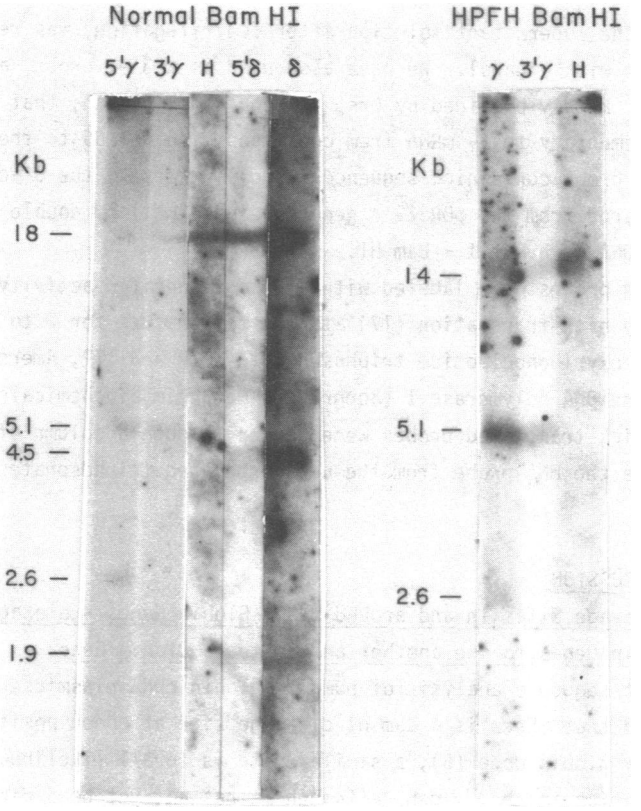

FIGURE 1. Gel blotting of normal DNA and HPFH DNA digested with Bam HI and fractionated by electrophoresis in slab gels of 0.7% agarose. Hybridization was carried out using the indicated nick translated plasmid probes (see Methods): 5'-γ cDNA & 3'-γ cDNA (from Eco RI + Hind III digested plasmid pHγGl [9]), total γ cDNA (plasmid JW151), H fragment DNA (Eco RI subclone of HβGl), 5'-δ DNA (from Bam HI + Pst I digested plasmid Hδ1 [Pst I subclone of HβGl]) and total δ DNA (plasmid Hδ1).

"first" γ gene, and the γ gene on the 3'-side as the "second" γ gene.

The fact that the 18 kb γ fragment in normal DNA also hybridizes strongly to 5'-δ globin gene probe (Figure 1) suggests that 18 kb is the minimum distance between the intragenic Bam HI sites of the second γ gene and of the δ gene. Because there is no direct experimental evidence enabling us to eliminate the possibility that this 18 kb band represents not one but two Bam HI fragments (one derived from the 3'-end of the second γ gene and the other from the 5'-end of the δ gene), it is possible that the γ-δ intergene dis-

tance may be 36 kb (or more) rather than 18 kb. This latter possibility however seems unlikely in view of the results obtained with HPFH DNA and with the use, as hybridization probe, of H fragment DNA that is derived from a fragment of DNA situated approximately 5 kb to the 5'-side of the δ gene in the clone HβG1 (10).

When HPFH DNA is cleaved with Bam HI, three fragments are obtained that hybridize to γ cDNA probes and have sizes of 14 kb, 5.1 kb and 2.6 kb (Figure 1). The 14 kb band hybridizes only to 3'-γ probe, the 2.6 kb band hybridizes only to 5'-γ probe and the 5.1 kb band hybridizes to both the 5'- and 3'-γ probes (5'-γ probe results not shown in Figure 1). The orientation of the HPFH Bam HI fragments in the direction of transcription is thus (5') 2.6 kb → 5.1 kb → 14 kb (3'). Therefore except for the 14 kb band, no other abnormality of Bam HI cleavage sites is detected in HPFH DNA compared to normal DNA. The 14 kb band, unlike the 18 kb band in normal DNA, does not hybridize to δ gene probe (data not shown). This result is consistent with the suggestion that the 18 kb Bam HI fragment in normal DNA contains both the 3'-end of one γ gene and the 5'-end of the δ gene and that in HPFH, the δ+β gene deletion (3,4) in the region of the δ gene includes approximately 4 kb of DNA to the 5'-side of the δ gene that are situated at the 3'-extremity of this 18 kb fragment (see section (g) below). This conclusion is reinforced by the results obtained using H fragment probe.

The H fragment probe hybridizes with both the 18 kb band of normal DNA, and the 14 kb band of HPFH DNA (Figure 1). This result is best explained if the normal γ-δ intergene distance is 18 kb, and the H fragment is located approximately 13 kb downstream from the intragenic Bam HI site of the second γ gene on the 18 kb Bam HI fragment of normal DNA. The 14 kb fragment of HPFH DNA would therefore similarly contain the 3'-end of the second γ gene and the H fragment DNA 13 kb downstream, but would lack most of the additional sequences, flanking and including the 5'-end of the δ gene, that are present in the normal 18 kb fragment. If, on the other hand, the γ-δ intergene distance were 36 kb (or more), one would have to postulate that the H probe and the γ probe hybridize to two separate 18 kb fragments in normal DNA and to two separate 14 kb fragments in HPFH DNA. Such an occurrence would require two separate and discontinuous deletion events in HPFH: one deletion starting approximately 14 kb downstream from the intragenic Bam HI site of the second γ gene and stopping somewhere upstream from the H fragment, and another deletion starting downstream from the preserved H fragment DNA sequence and extending through the δ and β globin genes to a point such that Bam HI diges-

tion of HPFH DNA would generate by coincidence a second fragment of DNA 14 kb in length that hybridizes to the H probe and that is separate from the 14 kb DNA fragment that contains the 3'-end of the second γ globin gene. Such a series of events and coincidences is extremely unlikely and the inter γ-δ gene distance is therefore approximately 16 kb (18 kb minus 2 kb of structural gene sequence). These results confirm those obtained by Fritsch et al.(22).

(b) Bgl II Cleavage Sites Flanking the γ Globin Genes

Digestion of both normal DNA and HPFH DNA with Bgl II yields a single prominent band, 13.5 kb in length, that hybridizes with the γ cDNA probe (Fig-

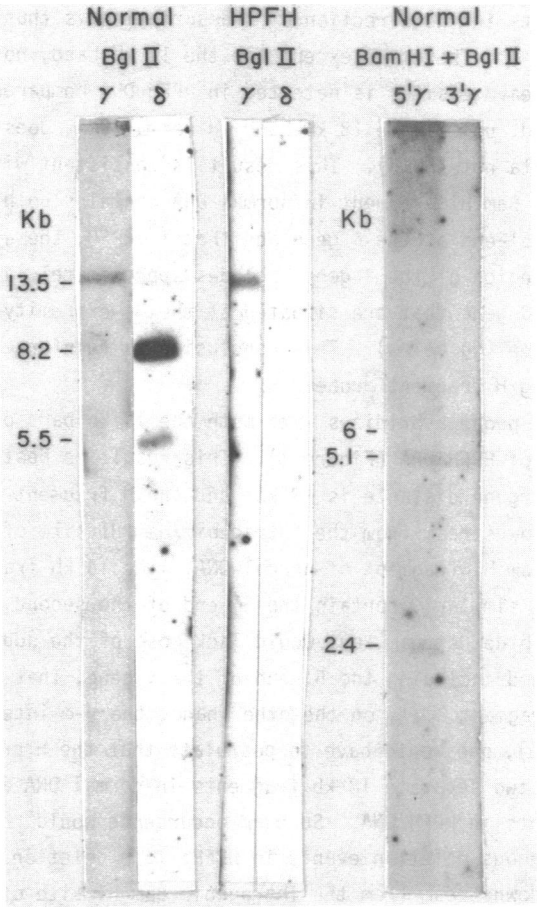

FIGURE 2. Gel blotting of normal DNA and HPFH DNA digested with Bgl II, and of normal DNA digested with Bam HI + Bgl II. Fractionation was by electrophoresis in slab gels of 0.7% agarose and hybridization probes were similar to those listed in Figure 1.

ure 2). To locate the Bgl II sites with respect to the Bam HI sites, DNA was doubly digested with Bam HI + Bgl II. The 5'- + 3'-γ 5.1 kb band generated by digestion with Bam HI alone is not cleaved by Bgl II; Bgl II therefore does not cut between the two γ genes. The 5'-γ 2.6 kb band obtained with Bam HI alone is slightly shortened by Bgl II to the size of 2.4 kb and the 3'-γ 18 kb band in Bam HI digests is shortened by additional digestion with Bgl II to the size of 6 kb (Figure 2). The 3'-γ 14 kb band in Bam HI digests of HPFH DNA is similarly shortened to 6 kb by Bgl II digestion (data not shown). The orientation of the three normal DNA fragments in Bam HI + Bgl II digests is thus (5') 2.4 kb → 5.1 kb → 6 kb (3'), and the sum of these three fragments is 13.5 kb, consistent with the fragment size obtained when the DNA is digested with Bgl II alone. From these results, Bgl II sites can be mapped to points 2.4 kb to the 5'-side of the intragenic Bam HI site of the first γ gene and 6 kb to the 3'-side of the intragenic Bam HI site of the second γ gene.

(c) Eco RI and Eco RI + Bgl II Cleavage Sites - Evidence for an Intervening Sequence in the γ Genes

From direct sequence analysis, it is known that the enzyme Eco RI cleaves both the $^{A}\gamma$ and $^{G}\gamma$ genes at codon positions 121 to 122 (6,7,9). After Eco RI digestion of normal as well as HPFH DNA, one obtains five bands, having sizes of 7.4 kb, 4.2 kb, 2.8 kb, 1.7 kb, and 0.7 kb respectively, that hybridize to γ globin cDNA probe (Figure 3). The presence of more than three bands indicates that Eco RI cuts at least once between the two linked γ globin genes. The Eco RI fragments must therefore include two derived from the 5'-ends and two derived from the 3'-ends of the two γ genes. The 7.4 and 2.8 kb bands hybridize only to 5'-γ cDNA probe (Figure 3). Since the distance between the two γ genes as determined from Bam HI digestion cannot be longer than 5.1 kb, the 7.4 kb fragment must be derived from the 5'-end of the first γ gene and the 2.8 kb fragment from the 5'-end of the second γ gene. The 1.7, 0.7 kb fragments hybridize only to 3'-γ probe. The 4.2 kb band hybridizes with the total γ cDNA probe but does not clearly hybridize with either the 5'-γ or 3'-γ half probes (Figure 3). A 4.2 kb band is also observed in normal DNA digested with Eco RI and hybridized to β probe (Figure 3d). The occurrence of a 4.2 kb band that hybridizes with γ cDNA is not easily explained by cross-hybridization of γ cDNA with a β globin gene fragment, because HPFH DNA in which the β gene as well as the δ gene are deleted also contains a similar 4.2 kb band after digestion with Eco RI (Figure 3b). Furthermore, the β globin gene fragment seen at the 4.2 kb position corresponds to that of the 3'-

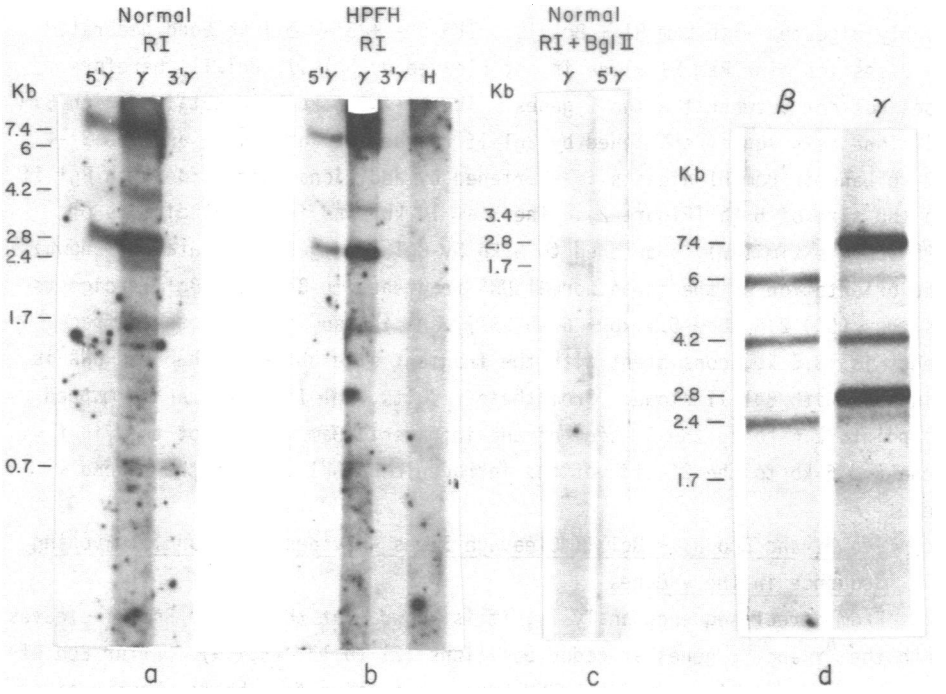

FIGURE 3. Gel blotting of human DNA fractionated by electrophoresis in slab gels of 1% agarose (a, b, c) or 0.7% agarose (d). Unless stated otherwise, the indicated hybridization probes were similar to those listed in Figure 1.
a) Normal DNA digested with Eco RI.
b) HPFH DNA digested with Eco RI.
c) Normal DNA digested with Eco RI + Bgl II.
d) Normal DNA digested with Eco RI. Hybridization probes were nick-translated β cDNA plasmid JW102 and γ cDNA plasmid JW151.

end of the β gene (23,24), and therefore consists mostly of 3'-noncoding sequences. The γ and β globin mRNAs display little homology in their 3'-untranslated sequences (6,7) and therefore the γ cDNA should give very little cross-hybridization with the Eco RI DNA fragment corresponding to the 3'-end of the β globin gene. Because the Bgl II + Bam HI digestion patterns of normal DNA are consistent with the presence of only two and not three γ globin genes, the 4.2 kb "γ" band may be derived not from an authentic γ gene but from an embryonic (ε) globin gene that has some sequence homology with the γ globin gene. In fact, the 4.2 kb band hybridizes less intensely with total γ cDNA probe than do the 7.4 kb and 2.8 kb γ bands; the lack of clearcut hybri-

dization of the 4.2 kb band with the 5'-γ and 3'-γ half-probes may reflect lack of perfect homology between γ and ε globin cDNA sequences. Although the $^{A}\gamma$ globin gene is known to have a Pst I site at codon positions 135 to 137 (7) that is lacking in the $^{G}\gamma$ gene (6), double digestion of DNA with Eco RI + Pst I did not lead to a significant mobility shift of either the 1.7 or 0.7 kb 3'-γ bands (data not shown); we are therefore unable to determine from which γ genes these two bands are derived.

As described above, Bam HI + Bgl II double digestion of human DNA yields a 2.4 kb DNA fragment that must be situated at the 5'-end of the first γ gene. In the absence of an intervening sequence in the γ gene, the Eco RI site at codon positions 121 to 122 of γ cDNA (6,7) should be located 60 bases downstream from the intragenic Bam HI site at codon positions 98 to 100, and an Eco RI + Bgl II double digest of the first γ gene should give a fragment similar in size to the 2.4 kb Bam HI + Bgl II 5'-γ fragment. Eco RI + Bgl II doubly digested normal DNA, when hybridized to 5'-γ cDNA probe, gives two bands having sizes of 3.4 kb and 2.8 kb (Figure 3C). Since Bgl II does not cut between the two γ genes, the 2.8 kb band must correspond to the 2.8 kb fragment derived from the 5'-end of the second γ gene by digestion of DNA with Eco RI alone (see above); the 3.4 kb band must therefore result from Bgl II cleavage of the 7.4 kb Eco RI band derived from the 5'-end of the first γ gene. This 3.4 kb fragment is 1 kb longer than expected. Since the same Bgl II site is located at the 5'-extremity of both the 2.4 kb Bam HI + Bgl II fragment and the 3.4 kb Eco RI + Bgl II fragment, an intervening sequence, approximately 1 kb in length, must be situated between the intragenic Bam HI site at codon positions 98 to 100 and the Eco RI site at codon positions 121 to 122 in the first γ globin gene. The total length of the first γ gene (the sum of the coding + intervening sequences) is therefore approximately 1.6 kb. From Hind III, Hind III + Bam HI, and Hind III + Eco RI digestion patterns of the second γ globin gene (presented below), it will be shown that the second γ gene also contains a 1 kb intervening sequence between its intragenic Bam HI and Eco RI sites. Analysis of a cloned human γ globin genomic DNA fragment has in fact identified an intervening sequence between codon positions 104 and 105 of this γ gene (25), at the same position where intervening sequences are located in the human β and δ globin genes (10).

(d) Hind III Cleavage Sites in and around the γ Globin Genes - Polymorphism for a Restriction Endonuclease Site within the Intervening Sequence of the γ Genes

No cleavage site for the enzyme Hind III has been identified in the se-

quence of γ globin cDNA (6,7,9,26). However, normal human DNA cleaved with Hind III gives five bands with sizes of 8 kb, 7.1 kb, 3.5 kb, 2.7 kb and 0.8 kb (Figure 4a). The 3.5 kb band hybridizes to both 5'-γ and 3'-γ cDNA probes, consistent with the interpretation that this fragment contains sequences complementary to a complete γ structural gene; this fragment is too small to be a fragment spanning the region between the two γ genes (see above). The 7.1 kb and 2.7 kb bands hybridize to 5'-γ cDNA probe only and the 0.8 kb band hybridizes only to 3'-γ cDNA probe (Figure 4a). The 7.1 kb band is probably derived from the 5'-end of the first γ gene; since the inter-γ gene distance is less than 5.1 kb (see above), the 7.1 kb band is too large to be derived from the 5'-end of the second γ gene. The 2.7 kb is then derived from the 5'-end of the second γ gene. There is thus an intragenic Hind III in both of the γ genes of this individual and this Hind III site must be located within the intervening sequence of the γ genes since the coding sequences of the γ genes do not contain Hind III sites. The presence of a 3.5 kb Hind III fragment as well as the 2.7 kb and 0.8 kb fragments (the sum of which equals 3.5 kb) suggests that the intragenic Hind III site is present in one but not both of the "second" γ genes in this sample of diploid DNA. In contrast, both of the "first" γ genes in the individual's diploid DNA probably contain an intragenic Hind III site. The 8 kb band seen in Figure 4a is significantly fainter than the 7.1 kb band and, unlike the 3.5 kb complete γ gene fragment, does not hybridize clearly with the 5'-γ or 3'-γ cDNA probes (Figure 4a). It is therefore possible that the 8 kb Hind III fragment in this individual is analogous to the 4.2 kb Eco RI fragment described above, and contains DNA that is derived from an embryonic ε globin gene.

In some DNA samples however, digestion with Hind III produces only 8 kb and 3.5 kb γ gene fragments (Figure 4c-1, 3, 4). Since the sum of the 7.1 kb + 0.8 kb partial fragments of the first γ gene is 7.9 kb, the 8.0 kb Hind III fragment in these individuals probably represents a complete γ gene fragment and the first γ gene of such individuals must lack an intragenic Hind III site. The presence of an intragenic Hind III site in the human γ genes therefore represents a polymorphism which in different individuals can be present or absent in either the $^{A}\gamma$ or $^{G}\gamma$ genes and can be manifested for either gene in the heterozygous or homozygous state.

When HPFH DNA is cut with Hind III, only four bands are obtained (8 kb, 7.1 kb, 3.5 kb, and 0.8 kb in length) that hybridize to the γ cDNA probe (Figure 4b). The absence of the 2.7 kb band indicates that in this HPFH individual, the γ gene DNA corresponding to the 3.5 kb fragment does not contain an

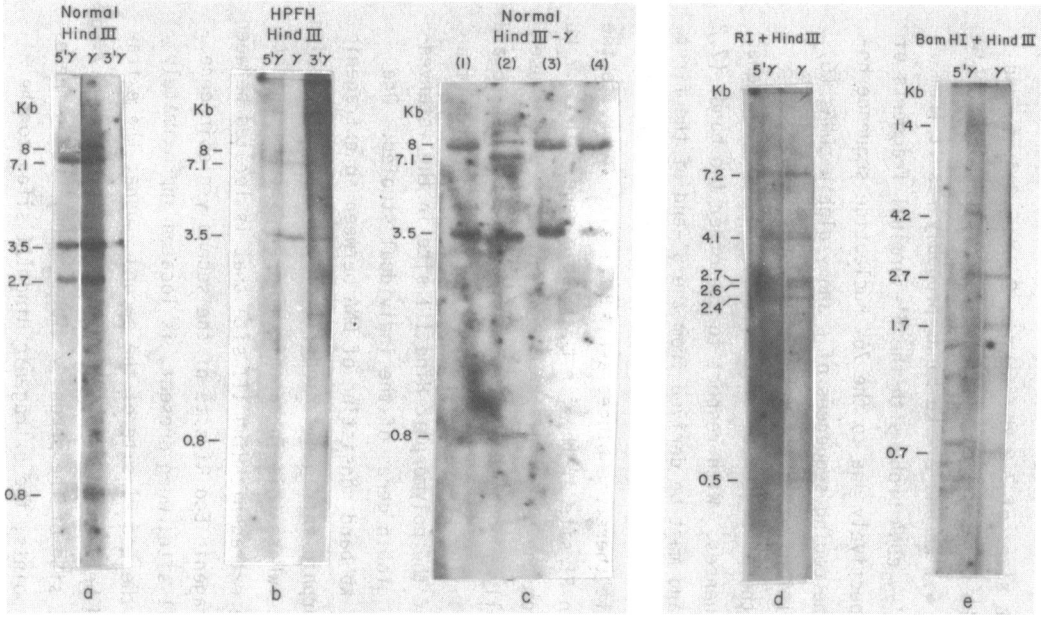

FIGURE 4. Gel blotting of human DNA fractionated by electrophoresis in slab gels of 1% agarose (a, d & e) or 1.2% agarose (b & c) and hybridized to the indicated probes (similar to those listed in Figure 1).

a) Normal DNA (same individual as in Figures 1 to 3), digested with Hind III.

b) HPFH DNA digested with Hind III.

c) DNA from four different non-HPFH individuals (listed below), digested with Hind III and hybridized with total γ cDNA probe.

1) DNA of δβ-thalassemic lymphoblastoid cell line (12), kindly provided by Drs. A. Bank and A. Greene.
2) Spleen DNA of Caucasian with autoimmune hemolytic anemia.
3) Peripheral blood DNA of black patient with sickle cell anemia.
4) Placental DNA of Asian infant with homozygous α thalassemia.

d) Normal DNA [same as in (a)] digested with Hind III + Eco RI.

e) Normal DNA [same as in (a)] digested with Hind III + Bam HI.

intragenic Hind III site in either chromosome. The absence of this intragenic site is unlikely to be associated with the activation of γ gene expression in HPFH, because a different non-HPFH DNA sample also lacks this site and gives the same Hind III digestion pattern as HPFH DNA (Figure 4c-2).

To map the Hind III sites around the γ genes in normal DNA, we carried out double digestion of DNA with Hind III + Eco RI: five bands hybridizing to γ cDNA were obtained, with sizes of 7.2 kb, 4.1 kb, 2.7 kb, 2.4 kb, and 0.5 kb. Except for the 0.5 kb band, all of the other bands hybridize to 5'-cDNA probe (Figure 4d). The 4.1 kb and 2.4 kb bands probably result from cross-hybridization of the 5'-γ cDNA probe with the 5'-terminal fragments of the β and δ globin genes respectively due to the 76% nucleotide sequence homology that exists between the coding sequences of β and γ globin cDNAs (6). Under the hybridization conditions utilized, we have frequently observed cross-hybridization between the 5'-sequences of the γ, β, and δ globin genes but not between their 3'-sequences. With respect to the other two bands (7.2 kb and 2.7 kb), the 7.2 kb band must be derived from the 5'-end of the first γ gene, since it is too large to be derived from the 5'-end of the second γ gene (see above). This result therefore places a Hind III site 7.2 kb to the 5'-side of the intragenic Eco RI site of the first γ globin gene. The 2.7 kb band is therefore derived from the 5'-end of the second γ gene. Upon closer inspection, this band is really a doublet consisting of two bands, 2.7 kb and 2.6 kb in length (Figure 4d). The presence of this doublet band can best be explained by the existence of the polymorphic Hind III site in the intervening sequence of one of the γ globin genes of the individual studied. One would thus obtain: 1) a 2.6 kb band consisting of DNA between an extragenic Hind III site and the polymorphic intragenic Hind III site of the intervening sequence, and 2) (in the case where this intragenic Hind III site is absent) a 2.7 kb band flanked by the extragenic Hind III site that is located between the two γ genes and the intragenic Eco RI site of the second γ globin gene. Thus, the intragenic Hind III site, when present, is located approximately 100 bases to the 5'-side of the Eco RI site of the second γ gene; the extragenic Hind III site situated between the two γ globin genes is located 2.7 kb to the 5'-side of the Eco RI site of the second γ globin gene. Using these Hind III sites as reference points, the extragenic Hind III sites to the 3'-side of the γ genes can be mapped to points 0.8 kb from the intragenic Hind III sites of both the first and second γ globin genes.

Since Hind III + Eco RI double digestion of the first γ gene produces a 7.2 kb 5'-fragment and a 0.8 kb 3'-fragment, and the sum of these two frag-

ments (8 kb) is equal to the size of the fragment obtained after digestion with Hind III alone, one can conclude that there is only one intragenic Eco RI site in the first γ gene. The same is also true of the second γ gene. The 5'-γ and 3'-γ Eco RI fragments of each γ gene (section c) are therefore contiguous without any intervening fragments generated by additional Eco RI sites possibly present in their intervening sequences.

The presence of an intragenic Hind III site indicates the existence of an intervening sequence between the Bam HI and Eco RI sites of the second γ globin gene as well as the first γ gene. To map its length, a Bam HI + Hind III double digest of DNA was carried out. Four bands were obtained (14 kb, 4.2 kb, 2.7 kb, and 1.7 kb in length) that hybridized with the 5'-γ probe (Figure 4e). The 14 kb and 4.2 kb bands probably correspond to DNA fragments of the δ and β globin genes respectively (10,24), revealed by cross-hybridization of the γ cDNA with δ and β gene sequences. The 2.7 kb band is most likely derived, by Bam HI cleavage alone, from the 5'-end of the first γ gene because the Hind III sites in this region lie outside of the two Bam HI sites (see above). The 1.7 kb band must therefore be derived from the 5'-end of the second γ globin gene. This 1.7 kb fragment must be flanked: 1) on its 3'-side, by a Bam HI site, because the closest Hind III site lies downstream in the intervening sequence or in the 3'-extragenic flanking sequence and, 2) on its 5'-end side, by an extragenic Hind III site because there is no Bam HI site between the two γ genes. As previously described, a 2.6 kb + 2.7 kb doublet fragment is obtained after Eco RI + Hind III double digestion of human DNA. Since both the 2.6 kb + 2.7 kb doublet and the 1.7 kb fragment (generated by Eco RI + Hind III and Bam HI + Hind III double digestions respectively) must be flanked on their 5'-ends by the same Hind III site, the 0.9 kb to 1 kb difference between these fragments indicates the presence of an intervening sequence of that length between the intragenic Bam HI and Eco RI sites of the second γ globin gene.

(e) Pst I Cleavage Sites in and around the γ Globin Genes - Orientation of the $^{G}\gamma$ and $^{A}\gamma$ Globin Genes

Cleavage of normal DNA with Pst I produces five major bands, with sizes of 12.5 kb, 5.1 kb, 4.0 kb, 2.8 kb and 0.9 kb respectively, that hybridize to γ cDNA probe (Figure 5a). Two additional bands are seen, 4.5 kb and 2.4 kb in length, that correspond to β and δ globin gene fragments (23,24) that cross-hybridize with the γ cDNA probe (Figure 5a). The 4.0 kb band hybridizes to 5'-γ cDNA probe only, and the 5.1 kb band hybridizes to both 5'- and 3'-γ cDNA probes (Figure 5b) (the 3'-γ cDNA probe in these experiments gave

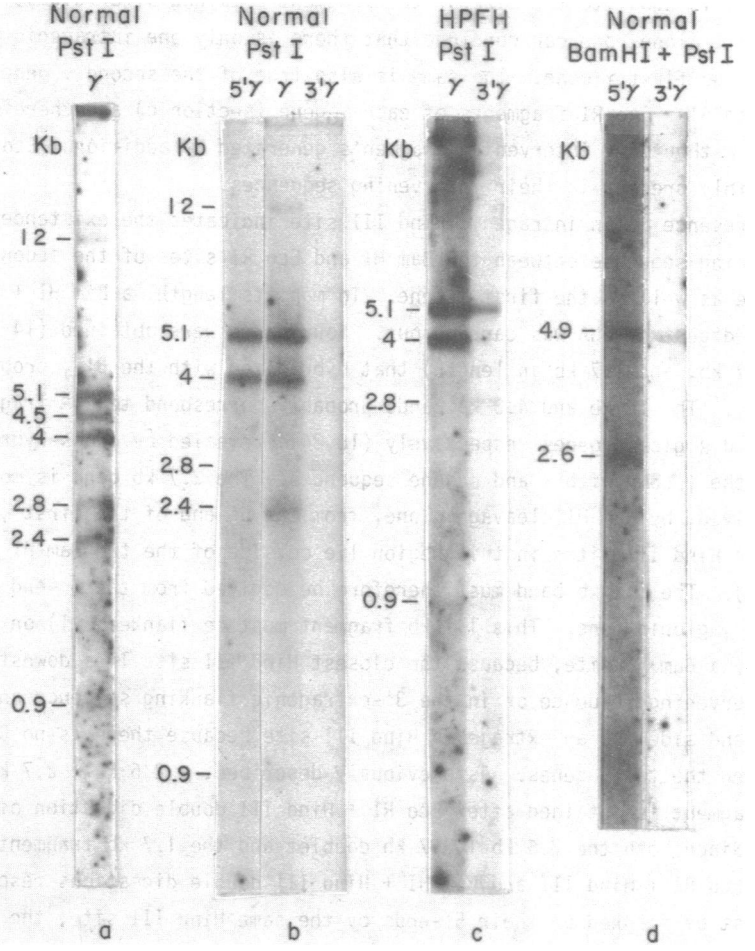

FIGURE 5. Gel blotting of human DNA fractionated by electrophoresis in slab gels of 0.7% agarose (a & d) or 1.0% agarose (b & c).
a) Normal DNA digested with Pst I. Total γ cDNA probe was from plasmid pHγG1 that contains sequences complementary to the 3'-untranslated sequence of γ mRNA (6).
b) and c) Normal DNA and HPFH DNA digested with Pst I. Total γ cDNA probe was from plasmid JW151 that lacks sequences complementary to the 3'-untranslated sequence of γ mRNA (6): the 2.8 kb fragment derived from the 3'-end of the A_γ gene is therefore not detected. The 3'-γ cDNA probe was from plasmid pRP10.
d) Normal DNA digested with Pst I + Bam HI. The 3'-γ and 5'-γ cDNA probes were from plasmid JW151 digested with Eco RI + Hha I.

only a weak positive signal). The 5.1 kb fragment therefore probably constitutes the DNA fragment spanning the two γ genes (Bam HI + Pst I digestion pat-

tern (Figure 5d) rules out the possibility that the 5.1 kb band contains an intact γ gene). The 2.8 kb and 0.9 kb bands hybridize only to 3'-γ cDNA probe (Figure 5b). The 12.5 band is faint and, like the 4.2 kb Eco RI fragment (and possibly one of the 8 kb Hind III fragments) is presumably derived from an embryonic ε globin gene fragment (see above); it hybridizes clearly only with the total γ cDNA probe but not with the 5'- or 3'-half probes (Figure 5b). If the 12.5 kb band is derived from the ε globin gene, then the other four bands are derived from the two γ globin genes. HPFH DNA cleaved with Pst I also produces these four bands (Figure 5c). Since the 4 kb Pst I band hybridizes only to 5'-γ probe and the 5.1 kb band is probably the fragment spanning the two γ genes, the four possible orientations of these Pst I fragments with respect to transcription (5'→3') are: 1) 4 kb → 0.9 kb → 5.1 kb → 2.8 kb; 2) 4 kb → 2.8 kb → 5.1 kb → 0.9 kb; 3) 4 kb → 5.1 kb → 2.8 kb → 0.9 kb; 4) 4 kb → 5.1 kb → 0.9 kb → 2.8 kb.

From nucleotide sequence analysis of the $^{G}\gamma$ and $^{A}\gamma$ cDNAs, the coding region of the $^{G}\gamma$ gene contains no Pst I site (6), whereas that of the $^{A}\gamma$ gene has one Pst I site at codon positions 135 to 137 (7). The generation of four bands by Pst I cleavage, including the 5.1 kb spanning fragment, indicates that the two γ genes contain at least three intragenic Pst I sites. The two γ globin genes must therefore each contain one Pst I site in their intervening sequences in addition to the site in the coding sequence of the $^{A}\gamma$ gene. In addition, the presence of the 5.1 kb spanning fragment indicates that these two Pst I sites are located at similar sites in the intervening sequences of both genes. For example, there are Bam HI sites at codon positions 98 to 100 in both the $^{A}\gamma$ and $^{G}\gamma$ genes that generate after digestion of total cellular DNA with Bam HI a 5.1 kb fragment that spans the two γ genes (see above); other restriction enzymes such as Xba I (see below), Hpa I and Hpa II (data not shown) that have symmetrically located cleavage sites in the two γ gene intervening sequences all produce a similarly sized 5.1 kb fragment that spans both γ genes. The 5.1 kb fragment also helps to map the probable location of the third intragenic Pst I site that must be situated at codon positions 135 to 137 in the $^{A}\gamma$ gene: it must be outside of the two Pst I sites of the intervening sequence, otherwise, a spanning fragment shorter than 5.1 kb would have been generated. This result therefore rules out potential orientation no. 1, (5') 4 kb → 0.9 kb → 5.1 kb → 2.8 kb (3'), which implies that the first γ gene contains two Pst I sites: one in the intervening sequence followed by another 0.9 kb downstream at codon positions 135 to 137, and that the second γ gene contains only one Pst I site, i.e., in its intervening se-

quence. Because the length of the γ gene is only approximately 1.6 kb, potential orientations no. 2 (5') 4 kb → 2.8 kb → 5.1 kb → 0.9 kb, and no. 3, (5') 4 kb → 5.1 kb → 2.8 kb → 0.9 kb (3'), specifying respectively that the first or the second γ gene contains two Pst I sites, 2.8 kb apart, can also be eliminated.

The one probable remaining orientation is therefore (5') 4 kb → 5.1 kb → 0.9 kb → 2.8 kb (3'). This orientation indicates that the first γ gene contains only one Pst I site and that the second γ gene contains two Pst I sites 0.9 kb apart. Since $^{G}\gamma$ is the γ gene that possesses one Pst I site and $^{A}\gamma$ the gene that has two Pst I sites, the orientation of the $^{G}\gamma$ and $^{A}\gamma$ genes with respect to transcription is therefore (5') $^{G}\gamma \rightarrow {}^{A}\gamma$ (3').

According to the orientation of the above Pst I fragments and by using the Pst I site in the $^{A}\gamma$ gene at codon positions 135 to 137 as a reference point, the other Pst I sites flanking the γ genes can be mapped as indicated in the summary map shown below in Figure 8. In mapping these Pst I sites, we have assumed that there were no undetectable Pst I fragments smaller than 0.5 kb and that, therefore, the four Pst I fragments detected are contiguous to one another. To confirm this point, double digestion of DNA with Bam HI + Pst I was carried out. A 4.9 kb fragment spanning the two γ genes was obtained (Figure 5d). The presence of this fragment, and the 5.1 kb fragment obtained after cleavage of DNA with Bam HI alone, indicates that the Pst I site in the $^{G}\gamma$ gene is situated no more than 0.2 kb to the 3'-side of the intragenic Bam HI site and must therefore be located in the intervening sequence of the $^{G}\gamma$ gene. Using this Pst I site as a reference point, a Pst I site can be mapped to an analogous position 5.1 kb downstream in the intervening sequence of the $^{A}\gamma$ gene (i.e., no more than 0.2 kb downstream from the $^{A}\gamma$ intragenic Bam HI site). Another intragenic Pst I site 0.9 kb downstream from this last Pst I site in the intervening sequence is therefore mapped to a point within the 3'-terminal coding region of the $^{A}\gamma$ gene at a position corresponding to codon positions 135 to 137.

(f) Xba I Cleavage Sites in and around the γ Globin Genes

Digestion of both normal and HPFH DNA with the enzyme Xba I yields DNA fragments with sizes of 7.5 kb, 5.1 kb and 3.6 kb that hybridize with γ cDNA (Figure 6). The 3.6 kb band hybridizes to 5'-γ cDNA probe only and the 7.5 kb band hybridizes primarily to 3'-γ probe; the 7.5 kb band also hybridizes faintly with the 5'-γ cDNA probe, as will be discussed below. The 5.1 kb band hybridizes to both 5'-γ and 3'-γ cDNA probes (Figure 6) (the 3'-γ cDNA hybridization gives a weak positive signal in these experiments also). Since

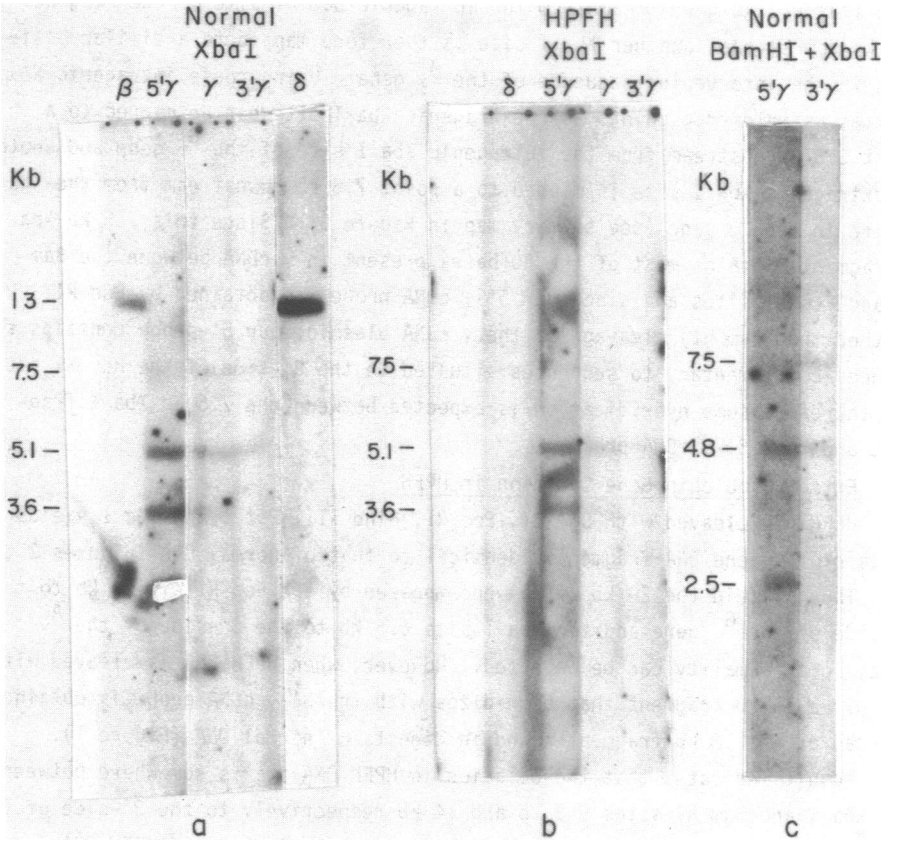

FIGURE 6. Gel blotting of normal DNA and HPFH DNA. Fractionation was by electrophoresis in slab gels of 0.7% agarose (Xba I single digests) or 1% agarose (Xba I + Bam HI double digest). Hybridization probes were similar to those listed in Figure 1.

Xba I does not cleave within the coding regions of $^{A}\gamma$ or $^{G}\gamma$ genes, the presence of this last fragment suggests that there is an Xba I site in the intervening sequences of both γ genes generating a 5.1 kb fragment spanning both γ genes. To locate the Xba I sites with respect to the intragenic Bam HI sites, Bam HI + Xba I double digestion of DNA was carried out. Three bands were obtained with sizes of 7.5 kb, 4.8 kb and 2.5 kb (Figure 6). The 4.8 kb band hybridizes to both 5'-γ and 3'-γ cDNA probes (Figure 6) and is therefore the fragment that spans the $^{G}\gamma$ and $^{A}\gamma$ genes. Thus the 5.1 Xba I fragment is reduced to 4.8 kb by Bam HI cleavage and this places an Xba I site no

more than 0.3 kb downstream from the intragenic Bam HI site of the $^{G}\gamma$ gene. 5.1 kb downstream, another Xba I site is therefore mapped to a similar position in the intervening sequence of the $^{A}\gamma$ gene. Using these intragenic Xba I sites as reference points, one extragenic Xba I site can be mapped to a point 3.6 kb upstream from the intragenic Xba I site of the $^{G}\gamma$ gene and another extragenic Xba I site is mapped to a point 7.5 kb downstream from the Xba I site in the $^{A}\gamma$ gene (see summary map in Figure 8). Since this 7.5 kb Xba I fragment contains most of the 60 bases present in γ cDNA between the Bam HI and Eco RI sites and since our 5'-γ cDNA probe was obtained by Eco RI (rather than Bam HI) cleavage of the γ cDNA plasmid, our 5'-probe contains sequences complementary to sequences situated to the 3'-side of the Bam HI site and therefore some hybridization is expected between the 7.5 kb Xba I fragment and this 5'-γ cDNA probe.

(g) Extent of Globin Gene Deletion in HPFH

HPFH DNA cleaved with Bgl II, Eco RI, Hind III, Pst I and Xba I yields a pattern of γ gene DNA fragments identical to that of normal DNA (Figures 2 to 6). Thus, within the 20 kb of DNA encompassed by the Eco RI site 6 kb to the 5'-side of the $^{G}\gamma$ gene and the Xba I site 6.5 kb to the 3'-side of the $^{A}\gamma$ gene, no abnormality can be detected. However, when HPFH DNA is cleaved with Bam HI, a 14 kb fragment that hybridizes with the 3'-γ cDNA probe is obtained instead of the 18 kb fragment found in digests of normal DNA (Figure 1). This result indicates that the deletion in HPFH DNA starts somewhere between the Xba I and Bam HI sites 6.5 kb and 14 kb respectively to the 3'-side of the $^{A}\gamma$ gene. Situated between these Xba I and Bam HI sites of HPFH DNA is the 0.5 kb of fragment H DNA (that is located to the 5'-side of the δ gene in normal DNA [10]) because a probe from fragment H hybridizes to the 14 kb Bam HI fragment of HPFH DNA (Figure 1). We have used the H fragment as a probe to map the extent and 5'-endpoint of the deletion in HPFH. Normal and HPFH DNA, digested with Eco RI and hybridized to the H probe, both yield a 7.2 kb fragment (Figure 7a). Since the H fragment is flanked on its 3'-side by an Eco RI site that has been mapped to a point approximately 4.5 kb upstream from the 5'-end of δ globin gene in normal DNA (10), the 3'-end of the 7.2 kb fragment must have this same Eco RI site; the 5'-end of this fragment is therefore 7.2 kb upstream,and both of these Eco RI sites are present in HPFH DNA. This result indicates that the deletion in HPFH starts at a point between the 3'-Eco RI site of the 7.2 kb fragment and the following 0.5 kb to 0.6 kb of DNA at which point the 3'-Bam HI site must be located in the 14 kb fragment of HPFH DNA (see map summary in Figure 8).

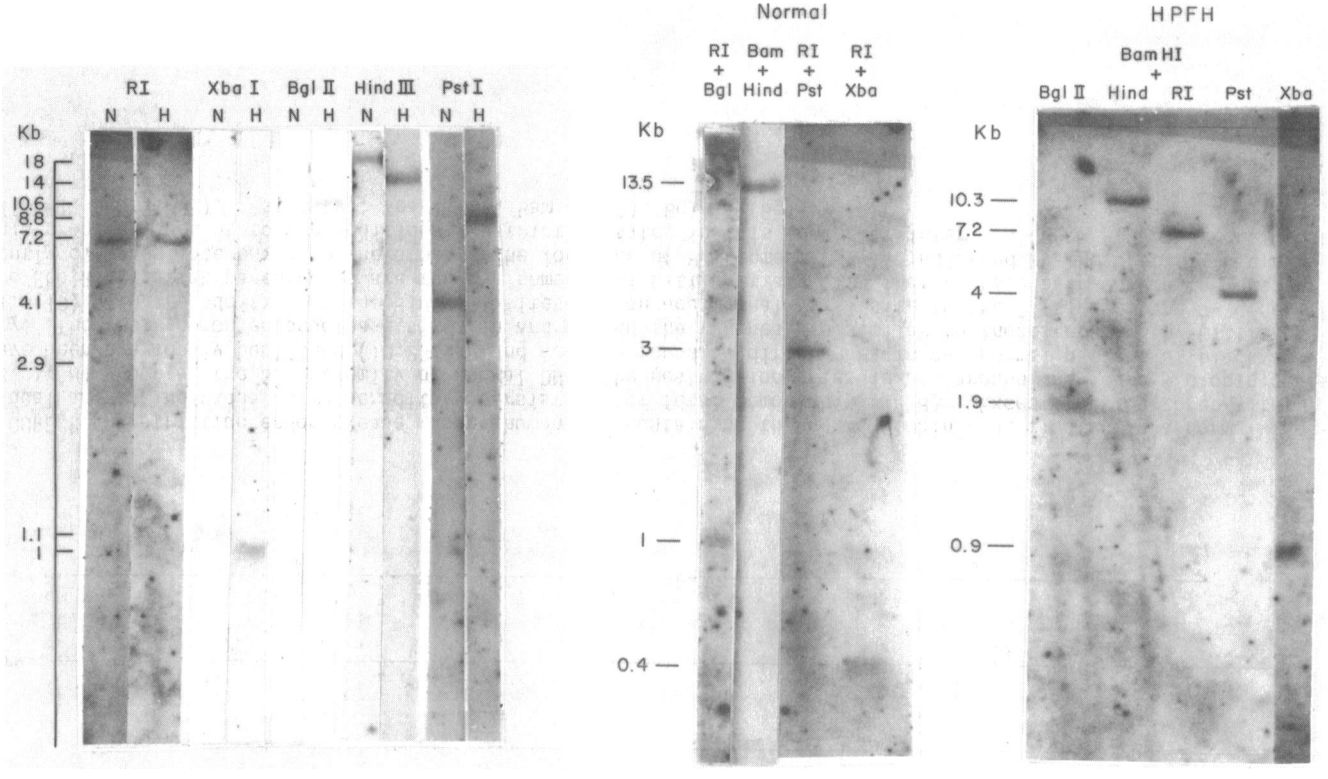

FIGURE 7. Gel blotting of human DNA using as hybridization probe nick-translated DNA from Eco RI subclone of HβGl that contains intergenic γ-δ DNA corresponding to fragment H (10) (see Figure 8). Fractionation was by electrophoresis in slab gels of 1% agarose.
a) Normal DNA (N) and HPFH DNA (H) singly digested with the indicated restriction endonucleases.
b) Normal DNA and HPFH DNA doubly digested with the indicated restriction endonucleases.

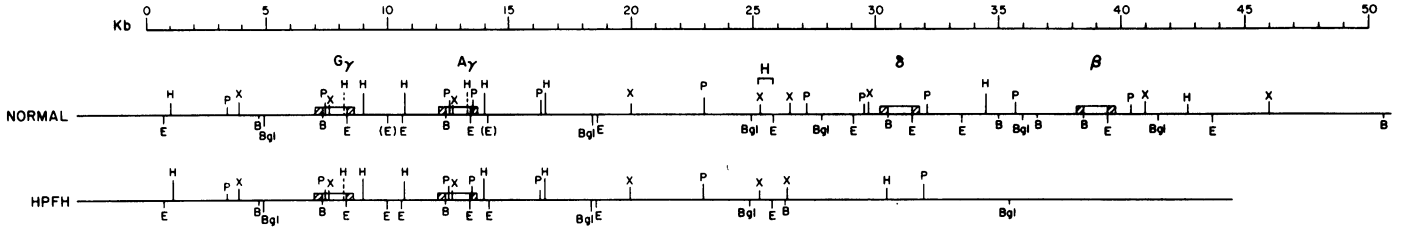

FIGURE 8. Restriction endonuclease map of human DNA containing the non-α globin gene cluster in a normal individual and an individual with hereditary persistence of fetal hemoglobin (HPFH). Except for the three sites (B, X, H) shown at the 3'-extremity of normal DNA, the restriction sites in and around the δ and β globin genes have been previously published (10,23,24) and we have found no differences in our normal patient of black ancestry. The restriction endonuclease sites in and around the γ genes are similar to those reported by Little et al. (19) with the addition of the Hind III sites and an additional Pst I site to the 3'-side of the $^{A}\gamma$ gene; the two Eco RI sites in parentheses are the assignments of Little et al. (19) that we have been unable to independently differentiate one from the other. The location of H fragment DNA is indicated by the bracket; use of this probe allowed us to map additional restriction sites in this region of inter γ-δ DNA. Abbreviations: Hind III (H), Eco RI (E), Pst I (P), Xba I (X), Bam HI (B), Bgl II (Bgl).

Normal DNA cleaved by Xba I and hybridized to H probe yields a 1.1 kb fragment, whereas HPFH DNA gives a 1.0 kb fragment. To map the Xba I sites relative to the Eco RI site at the 3'-end of H fragment (or to the Bam HI site 0.6 kb downstream in HPFH DNA), Eco RI + Xba I double digestion of normal DNA, and Bam HI + Xba I of HPFH DNA were carried out (Figure 7b). Bam HI + Xba I cleavage of HPFH DNA produces a 0.9 kb fragment, which therefore places an Xba I site 0.9 kb to the 5'-side of the Bam HI site in HPFH. Since this Xba site maps to a point in the vicinity of the 0.5 kb H fragment, this Xba I site must also be present in normal DNA, and cleavage of normal DNA with Eco RI + Xba I in this area should yield a fragment of DNA 0.4 kb in length, having at its 3'-end the 3'-Eco RI site of the H fragment (Figure 8). Eco RI + Xba I double digestion of normal DNA in fact produces a poorly defined band 0.4 kb to 0.5 kb in length (Figure 7b); this size range is at the lower limit of resolution in agarose gels. Using this Xba I site as the reference point, the other Xba I sites in normal and HPFH DNA can be mapped to points 1.1 kb and 1.0 kb downstream. The 3'-Xba I site is shifted compared to normal in HPFH DNA because it is located in DNA to the 3'-side of the deletion in HPFH.

Other restriction endonucleases produce fragments from HPFH DNA that differ in size from those of normal DNA. For example, Bgl II, Hind III and Pst I generate from HPFH DNA fragments with sizes of 10.6 kb, 14 kb, and 8.8 kb respectively, whereas the corresponding fragments from normal DNA have lengths of 2.9 kb, 18 kb, and 4.1 kb respectively (Figure 7a). To map these sites relative to the 3'-Eco RI site of the H fragment or the 3'-Bam HI site of the 14 kb HPFH fragment, Bam HI + Bgl II (or + Hind III or + Pst I) double digestions of HPFH DNA were carried out as well as Eco RI + Bgl II (or + Pst) and Bam HI + Hind III double digestions of normal DNA (Figure 7b). Knowing the lengths of single and double digestion fragments and the location of the Eco RI and Bam HI reference sites, these restriction endonuclease sites around the H fragment can be mapped in normal and HPFH DNA (Figure 8). In Figure 8, the restriction endonuclease sites in the DNA to the 3'-side of the deletion in HPFH are mapped (in relation to the 3'-Eco RI site of the H fragment) over a distance of 10 kb,at the 3'-end of which is present a Bgl II site that is not mapped in normal DNA.

In an attempt to determine the endpoint of the deletion in relation to normal DNA beyond the β globin gene, we have used a probe derived from the Eco RI D fragment of clone HβGl (10) that consists of approximately 3 kb of DNA to the 3'-side of the intragenic Eco RI site of the β globin gene. HPFH

DNA cleaved with Xba I or Bgl II, when hybridized to a probe of nick-translated D fragment, yields no visible bands (data not shown). The deletion in HPFH therefore stretches from its 5'-endpoint near the H fragment for at least 16 kb, and extends beyond a point 3 kb from the 3'-end of the β globin gene.

(h) Summary and Conclusions

The differentially activated and expressed human non-α globin genes ($^{G}\gamma$, $^{A}\gamma$, δ and β) are known from genetic analysis (1,2) to be linked on the same chromosome and work with somatic cell hybrids has further localized these genes to human chromosome no. 11 (27). The restriction endonuclease mapping presented in this report confirms the linkage of these genes to within a 35 kb stretch of chromosomal DNA. The orientation of the non-α globin genes with respect to the direction of transcription is (5') $^{G}\gamma\rightarrow{}^{A}\gamma\rightarrow\delta\rightarrow\beta$ (3') and the distances between these genes are 3.5 kb ($^{G}\gamma\rightarrow{}^{A}\gamma$) and 16 kb ($^{A}\gamma\rightarrow\delta$) in agreement with results reported by others (19,22); the distance between the δ and β globin genes is approximately 5 to 6.5 kb (10,23,24). Despite their close linkage however, the expression of these genes is not synchronous: the $^{G}\gamma$ and $^{A}\gamma$ genes are expressed maximally in fetal life, then after birth their expression is reduced to the level of only 1 to 2% of that of the adult β (and δ) globin genes. The "suppression" of fetal γ globin gene expression in adult life is not the result of γ globin gene deletion or rearrangement, because both γ genes are present in adult tissues and their orientation with respect to δ and β globin genes is unchanged with respect to that observed in fetal tissues (unpublished observations).

We have also studied the non-α globin gene loci in a case of homozygous HPFH, a genetic disorder in which neither the δ nor the β globin chains are synthesized in adult reticulocytes (due to deletion of δ and β globin genes [3,4]) but in which the fetal γ globin genes continue to be expressed at a very high level thus compensating for the δ and β globin chain deficiency. Restriction endonuclease mapping of HPFH DNA shows no sequence abnormality in or near the γ gene loci. Both γ globin genes are present and, as in normal DNA, both γ genes contain intervening sequences 1 kb in length with identical restriction endonuclease sites. Although the intervening sequence in $^{A}\gamma$ genes of this HPFH individual lacks a Hind III site, the absence of this site is probably not related to the persistent expression of γ genes in HPFH, since other non-HPFH DNA samples also lack this site which appears to be variably present or absent in either $^{A}\gamma$ or $^{G}\gamma$ gene in different individuals.

12.5 kb to the 3'-side of the $^{A}\gamma$ gene in HPFH, the start point of the δ

+ β globin gene deletion has been identified: from this point there have been deleted at least 16 kb of DNA that are present in normal DNA, including the entire δ and β globin genes. The deletions of the δ and β globin genes in themselves do not seem to constitute the basis for the persistent γ gene expression in HPFH, because in another type of genetic disorder, δβ-thalassemia, in which the majority of the δ gene as well as the entire β globin gene are deleted (Ref. 12,22 and unpublished observations), γ globin gene expression does not adequately compensate for the absence of δ and β globin chain synthesis (1). One definite difference between these specific cases of δβ-thalassemia (12) and HPFH (4) is in the starting points of the two deletions. In this case of δβ-thalassemia, the deletion starts in the middle of the δ globin gene to the 3'-side of the intragenic Bam HI site and continues beyond the β globin gene (22, 28 and unpublished observations) whereas in HPFH, the deletion starts at about 4 kb to the 5'-side of the 5'-end of the δ globin gene and extends beyond the β globin gene. It is thus possible that within the 4 kb of DNA situated between the starting point of the HPFH deletion and the 5'-end of the δ globin gene, there may exist sequences that control γ globin gene expression and that the deletion of the sequences in HPFH (but not in the case of δβ-thalassemia) results in continued expression of the γ genes in adult life in all of the HPFH individual's erythroid cells. Another possibility is that the end points of the deletions in HPFH and δβ-thalassemia may differ on the 3'-side of the β globin genes and that sequences in this area may be responsible for regulating γ gene expression and the differential extent of fetal hemoglobin synthesis in HPFH and δβ-thalassemia. In fact, one type of δβ-thalassemia has been identified (22) in which the $^{A}\gamma$ gene and all of the DNA between the $^{A}\gamma$ and δ globin genes has been deleted. Absence of the 4 kb of DNA adjacent to the 5'-end of the δ gene in this individual did not result in a HPFH phenotype.

ACKNOWLEDGEMENTS

We thank Dr. T. Maniatis and his co-workers for helpful discussions and for providing us with the human δ + β genomic clone HβG1 and its various subclones. We thank Dr. S. Malcolm for providing us with the γ cDNA plasmid clone pHγG1 and Drs. R. Poon and Y.W. Kan for the γ cDNA clone pRP10. E. Coupal, E. Hatton and V. Vellucci provided skilled technical assistance. This work was supported in part by grants from the National Science Foundation (PCM77-15963 [D.T.]) and the National Institutes of Health (HL20922 and AM19482 [B.G.F.], and GM20124 [P.A.B.]).

REFERENCES

1. Weatherall, D.J. and Clegg, J.B. (1972) The Thalassaemia Syndromes. 2nd ed. Blackwell Scientific Publications, Oxford.

2. Bunn, H.F., Forget, B.G., and Ranney, H.M. (1977) Human Hemoglobins. W.B. Saunders, Co., Philadelphia.
3. Kan, Y.W., Holland, J.P., Dozy, A.M., Charache, S., & Kazazian, H.H. (1975) Nature 258, 162-163.
4. Forget, B.G., Hillman, D.G., Lazarus, H., Barell, E.F., Benz, E.J. Jr., Caskey, C.T., Huisman, T.H.J., Schroeder, W.A., and Housman, D. (1976) Cell 7, 323-329.
5. Davidson, E.H. and Britten, R.J. (1973) Quarterly Rev. Biology 48, 565-620.
6. Forget, B.G., Cavallesco, C., deRiel, J.K., Spritz, R.A., Choudary, P.V., Wilson, J.T., Wilson, L.B., Reddy, V.B., and Weissman, S.M. (1979) In "Eukaryotic Gene Regulation." R. Axel, T. Maniatis, and C.F. Fox (eds.) ICN-UCLA Symposia (in press).
7. Poon, R., Kan, Y.W., and Boyer, H.W. (1978) Nucleic Acids Res. 5, 4625-4630.
8. Wilson, J.T., Wilson, L.B., deRiel, J.K., Villa-Komaroff, L., Efstratiadis, A., Forget, B.G., and Weissman, S.M. (1978) Nucleic Acids Res. 5, 563-581.
9. Little, P., Curtis, P., Coutelle, C., van den Berg, J., Dalgleish, R., Malcolm, S., Courtney, M., Westaway, D., and Williamson, R. (1978) Nature 273, 640-643.
10. Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G., and Maniatis, T. (1978) Cell 15, 1157-1174.
11. Southern, E. (1975) J. Mol. Biol. 98, 503-517.
12. Mears, J.G., Ramirez, F., Leibowitz, D., Nakamura, F., Blood, A., Konotey-Ahulu, F., and Bank, A. (1978) Proc. Nat. Acad. Sci. (USA) 75, 1222-1226.
13. Blin, N. and Stafford, D.W. (1976) Nucleic Acids Res. 3, 2302-2308.
14. McDonnell, M.W., Simon, M.N., and Studier, F.W. (1977) J. Mol. Biol. 110, 119-146.
15. Maxam, A.M. and Gilbert, W. (1977) Proc. Nat. Acad. Sci. (USA) 74, 560-564.
16. Cory, S. and Adams, J.M. (1977) Cell 11, 795-805.
17. Maniatis, T., Jeffrey, A., and Kleid, D.G. (1975) Proc. Nat. Acad. Sci. (USA) 72, 1184-1188.
18. Jeffreys, A.J. and Flavell, R.A. (1977) Cell 12, 429-439.
19. Little, P.F.R., Flavell, R.A., Kooter, J.M., Annison, G., and Williamson, R. (1979) Nature 278, 227-231.
20. Ramirez, F., O'Donnell, J.V., Natta, C., and Bank, A. (1976) J. Clin. Invest. 58, 1475-1481.
21. Old, J., Clegg, J.B., Weatherall, D.J., Ottolenghi, S., Comi, P., Giglioni, B., Mitchell, J., Tolstoshev, P., and Williamson, R. (1976) Cell 8, 13-18.
22. Fritsch, E.F., Lawn, R.M., and Maniatis, T. (1979) Nature (in press).
23. Mears, J.G., Ramirez, F., Leibowitz, D., and Bank, A. (1978) Cell 15, 15-23.
24. Flavell, R.A., Kooter, J.M., DeBoer, E., Little, P.F.R., and Williamson, R. (1978) Cell 15, 25-41.
25. Smithies, O., Blechl, A.E., Denniston-Thompson, K., Newell, N., Richards, J.E., Slightom, J.L., Tucker, P.W., and Blattner, F.R. (1978) Science 202, 1284-1289.
26. Chang, J.C., Poon, R., Neumann, K., and Kan, Y.W. (1978) Nucleic Acids Res. 5, 3515-3522.
27. Deisseroth, A., Nienhuis, A., Lawrence, J., Giles, R., Turner, P., and Ruddle, F. (1978) Proc. Nat. Acad. Sci. (USA) 75, 1456-1460.
28. Ottolenghi, S., Giglioni, B., Comi, P., Gianni, A.M., Polli, E., Acquaye, C.T.A., Oldham, J.H., and Masera, G. (1979) Nature 278, 654-657.