Structure of the human $^{G}\gamma$ - $^{A}\gamma$ - δ -globin gene locus

(genomic blotting/partial digests/restriction map/intergene distances)

R. Bernards*, P. F. R. Little[†], G. Annison[†], R. Williamson[†], and R. A. Flavell*

*Universiteit van Amsterdam, Laboratorium voor Biochemie, Jan Swammerdam Instituut, Eerste Constantijn Huygensstraat 20, Amsterdam, The Netherlands; and †Department of Biochemistry, St. Mary's Hospital Medical School, University of London, London W2 1PG, England

Communicated by Donald D. Brown, June 15, 1979

ABSTRACT We have constructed a physical map of the human ${}^{C}\gamma$, ${}^{A}\gamma$, δ , and β -globin genes. The previously described maps of the fetal and adult β -like globin genes have been linked to one another by identification of a DNA fragment, generated by BamHI, that contains part of each of the ${}^{A}\gamma$ - and δ -globin genes. The map obtained, which spans more than 40 kilobases, shows the following intergene distances: between ${}^{C}\gamma$ and ${}^{A}\gamma$, 3500 base pairs; between ${}^{A}\gamma$ and δ , 13,500 base pairs; and between δ and β , 5500 base pairs. All genes are transcribed from the same DNA strand.

The human β -like $(\beta$ -, δ -, and γ -) globin genes show a coordinated developmental program of sequential activation (1). In very young embryos (up to 3 months) the hemoglobin molecule functions with embryonic ϵ chains. These are replaced within the first 2 or 3 months of fetal life by the γ -globin chains (2). There are two nonallelic γ -globins, $^A\gamma$ and $^G\gamma$, with either alanine or glutamic acid at position 136 of the 146-amino-acid γ -globin protein chain (3). There is a gradual switch from fetal to adult globin chain synthesis, which is initiated at about 36 weeks after conception and is only completed some 6 months after birth (4). In adults, hemoglobin contains mostly β -globin chains, with 2–4% of the molecules containing the δ -globin chains

In two previous reports (5, 6) we have described the separate physical "maps" of the linked $^{A}\gamma$ - and $^{G}\gamma$ -globin genes and the linked β - and δ -globin genes. Our method of analysis involves cutting human DNA with a variety of restriction endonucleases, singly and in double digests, separation of DNA fragments by agarose gel electrophoresis, transfer of DNA to nitrocellulose filters (7), and identification of DNA fragments containing globin genes by hybridization to ³²P-labeled recombinant human globin plasmids. These contain double-stranded cDNA fragments specific for β - or γ -globin genes. We then construct a map of relative positions of restriction sites; this has made possible the demonstration of intervening sequences in all four genes, the determination of the gene orientation, and the determination of the β - δ and ${}^{\rm G}\gamma$ - ${}^{\rm A}\gamma$ intergene distances. In this communication we demonstrate the linkage of the two maps already derived and describe the physical structure of the human $^{G}\gamma$ - $^{A}\gamma$ - δ - β -globin gene locus.

MATERIALS AND METHODS

Transfer Hybridization. Details of DNA isolation from hematologically normal human placentas or the spleen of a homozygous Lepore patient and restriction enzyme digestion have been reported (5, 6).

For partial restriction endonuclease digests, 200 μ g of DNA per ml was digested with 100 units of the appropriate enzyme per ml and samples were isolated after 10, 20, and 40 min.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

For total DNA analysis, 15 µg of digested double-stranded DNA was run on a 0.7%, $20 \times 20 \times 1$ cm horizontal agarose gel at a voltage gradient of 1 V/cm for 18 hr. For RPC-5-fractionated DNA, between 1 and 2 µg of DNA from each fraction was run on a 1%, $20 \times 20 \times 0.4$ cm horizontal agarose gel at a voltage gradient of 0.5 V/cm for 17 hr. DNA was transferred to nitrocellulose filters (7) and DNA fragments containing β -, δ -, or γ -globin genes were detected by hybridization with the appropriate ³²P-labeled recombinant probe. Conditions of hybridization were as described (8); two probes were used, pH β G1 and pH γ G1 (9). These plasmids are both recombinant pCRI plasmids containing approximately 540 and 500 base pairs of double-stranded cDNA and cDNA, respectively. Both plasmids were grown in the London laboratory under CII conditions in a nondisabled host, as advised by the United Kingdom Genetic Manipulation Advisory Committee.

Whole pH β G1 or pH γ G1 DNA was labeled by nick translation. For some experiments, probes specific for the 3' or 5' side of the *Bam*HI site coding for amino acids 98–100 of the β -globin chain (6) were used.

Procedures for posthybridization washes and conditions of autoradiography were as described (5, 6, 8). Sizes of globin gene-containing fragments were calculated for their mobility relative to 32 P-5'-end-labeled markers of phage λ DNA [46 kilobases (kb)], phage λ DNA digested with *Eco*RI (20.5, 7.1, 5.6, 5.2, 4.5, and 3.2 kb), phage ϕ 29 DNA digested with *Eco*RI (9.2, 5.8, 2.0, 0.96, and 0.67 kb), and phage ϕ X174 RF DNA digested with *Taq* I (2.906 and 1.172 base pairs and smaller). For the *Kpn* I fragment described in the text, internal markers of λ DNA and λ DNA digested with *Eco*RI were used.

RPC-5 Chromatography. BamHI-digested human DNA (2.5 mg), dissolved in 1.0 M NaOAc/50 mM Tris-HCl/1 mM EDTA, pH 7.5, was loaded onto a 32-ml column of RPC-5 (Miles batch 7), and DNA was eluted with a 200-ml gradient of 1.5–1.7 M NaOAc/50 mM Tris-HCl/1 mM EDTA, pH 7.5. Operating pressure was 200–400 lbs./inch² (1 lb./inch² = 6.9 \times 10³ pascals), with a Milton Roy Minipump. Fractions containing DNA were precipitated with ethanol and taken up in a constant volume of 10 mM Tris-HCl/1 mM EDTA, pH 7.5.

RESULTS

In our previous work we observed that BamHI, which cuts both the δ - and $^{\Lambda}\gamma$ -globin genes at the position coding for amino acids 98–100, generated a 15-kb fragment containing the 5' end of the δ gene and a 15-kb fragment containing the 3' end of the $^{\Lambda}\gamma$ gene. Genetic evidence and direct restriction mapping have shown that the gene order in this locus is $^{G}\gamma$, $^{\Lambda}\gamma$, δ , and β and that each gene is transcribed from the same DNA strand (1, 5, 6). Thus it seemed possible that the 15-kb BamHI fragment linked the $^{\Lambda}\gamma$ - and $^{\delta}$ -globin genes. We would, on this basis,

Abbreviation: kb, kilobases.

predict that at its termini this 15-kb fragment would contain 190 base pairs of mRNA γ -coding sequence at one end (from the position coding for amino acid 101 to the 3′ end) and at its other end about 300 base pairs of mRNA δ -coding sequence (from the 5′ end to the position coding for amino acid 99). The alternative hypothesis is that $Bam{\rm HI}$ generates two 15-kb fragments, one containing the δ -globin gene fragment and one the $^{\rm A}\gamma$ -globin fragment. To test this, we used RPC-5 chromatography to attempt to separate these putative two fragments. Fig. 1 shows that, whereas the various smaller $Bam{\rm HI}$ γ , δ , and β fragments are separated by RPC-5, the 15-kb fragment hybridizing to the γ -globin and δ -globin probes cochromatographs.

If there is a single 15-kb BamHI fragment bridging the γ and δ -globin gene regions, then cleavage by another restriction enzyme at a single site within this fragment should generate two double-digest fragments, one of which contains part of the $^{\rm A}\gamma$ -globin gene and one of which contains part of the δ -globin gene. The lengths of the double-digest fragments should add up to 15 kb. We have screened a number of restriction endonucleases for sites in the 15-kb BamHI fragment. Of these, only Hpa I cleaved only once in this region. A HpaI/BamHI double digest generated a 1.3-kb 5' δ-globin gene fragment and a 14-kb fragment containing the 3' region of the $^{\rm A}\gamma$ -globin gene (Fig. 2). In this experiment, two additional bands (marked ◆) are seen in the BamHI and BamHI/Hpa I lanes. These are caused by picogram amounts of pH\gammaG1 plasmid DNA contaminating the human DNA used for this experiment (see, e.g., Figs. 1 and 4, where BamHI digests do not exhibit these bands: numerous published BamHI double and single digests hybridized with γ probes do not show these bands). In other *Hpa* I/BamHI double digests hybridized with the γ probe, these contaminant bands are missing and only the 14-, 5-, 3.8-, and 2.7-kb bands can be seen. The presence of a single *Hpa* I site in both γ and $\delta\beta$ experiments is consistent with a single 15-kb BamHI fragment containing parts of the δ - and $\bar{\Lambda}\gamma$ -globin genes. Two restriction endonucleases, BstEII and Kpn I, did not cleave the 15-kb BamHI fragment.

An alternative approach to establish the linkage of the γ - and δ -globin genes by the 15-kb BamHI fragment is to localize the sites of restriction endonuclease cleavage for enzymes that cut more than once in the 15-kb fragment. The position of the sites for each enzyme should be identical when located with either

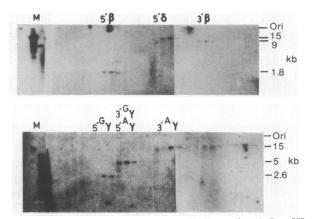


FIG. 1. Fractionation of human DNA digested with $Bam{\rm HI}$ on RPC-5. Aliquots of DNA eluted from the RPC-5 column and precipitated with ethanol were run on agarose gels, blotted, and hybridized with γ - and β -specific probes separately (Lower and Upper, respectively). Note that the 15-kb γ - and δ -globin $Bam{\rm HI}$ fragments comigrate on RPC-5. M represents marker DNA tracks and each fragment is identified. Ori is the origin. The $3'Bam{\rm HI}$ δ fragment is not seen with the pH β G1 probe (6).

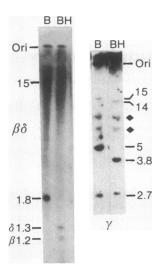


FIG. 2. Location of an Hpa I site in the 15-kb BamHI fragment. Human DNA cut with BamHI (B) was further digested to completion with Hpa I (BH) and 15- μ g samples were run on agarose gels, blotted, and hybridized with either $^{32}\text{P-labeled pH}\gamma\text{G1}$ (γ) DNA or with the 5'-specific pH β G1-derived DNA probe ($\beta\delta$). \blacklozenge , Plasmid contaminants.

 γ - or δ -globin probes. To determine the position of the cleavage sites for two such enzymes, Bcl I and Bgl II, in the 15-kb BamHI fragment, we first cleaved human DNA with BamHI. The BamHI fragments generated from the γ -, δ -, and β -globin genes are shown in Fig. 3. We then made partial digests with Bcl I or Bgl II on the BamHI-cut DNA. γ -Globin gene fragments were detected with 32 P-labeled pH γ G1 DNA and the 5' β and δ fragments were detected with the 5' β probe. Any partial fragment(s) detected between 15 kb and the next largest BamHI fragment (5 kb for the γ genes and 1.8 kb for the $\delta\beta$ genes) must be derived from the 15-kb BamHI fragment.

The 15-kb BamHI fragment was cleaved by Bcl I (Fig. 3) to give a 2.1-kb double-digest fragment containing the 5' region of the δ -globin gene. Partial digestion of the 15-kb BamHI fragment with Bcl I yielded an additional band at 9.2 kb that hybridized with probes for the δ -globin gene. This partial digest fragment, therefore, predicts a Bcl I site 5.8 kb (15 – 9.2 kb) from the " γ " end of the BamHI fragment. A total digest of Bcl I and BamHI exhibited a 6-kb band containing the 3' regions of the $^{\Lambda}\gamma$ -globin gene (Fig. 3 and Table 1). This site, close to the $^{\Lambda}\gamma$ gene, is therefore at the position predicted from the partial fragment containing the δ gene, assuming that the 15-kb BamHI fragment links the two genes.

Similarly, partial digestion of the 15-kb BamHI fragment with Bcl I hybridized with γ -gene probes showed a 13-kb fragment. If we assume that the 15-kb BamHI fragment contains both the $^{\Lambda}\gamma$ - and δ -globin gene regions, then this partial fragment predicts a Bcl I site 2 kb (15 - 13 kb) from the BamHI site in the δ gene; this is the fragment of 2.1 kb found in digests of Bcl I/BamHI hybridized with a δ -globin probe. These data are consistent with a single 15-kb BamHI fragment containing the 3' $^{\Lambda}\gamma$ -globin gene sequence and the 5' δ -globin gene sequence.

We have also performed Bgl II partial digests on the 15-kb BamHI fragment. When probes for the δ -globin gene were used, a 3.0-kb BamHI total digest fragment was seen, together with partial fragments of 5.9 and 9.2 kb (Fig. 4 and Table 2). The 15-kb BamHI fragment, therefore, contains three Bgl II sites 3.0, 5.9, and 9.2 kb from the BamHI site in the δ -globin gene. The 9.2-kb partial fragment detected by δ probes predicts that the Bgl II site closest to the $^{\Lambda}\gamma$ -globin gene will be at a distance of about 6 kb from the BamHI site in the γ gene (15 – 9.2 kb) and that two other sites should be detectable in partial digests. We have previously shown that a total digest of Bgl II/BamHI, when hybridized with γ probes, exhibits just such a 6-kb fragment containing the 3' regions of the $^{\Lambda}\gamma$ gene (5).

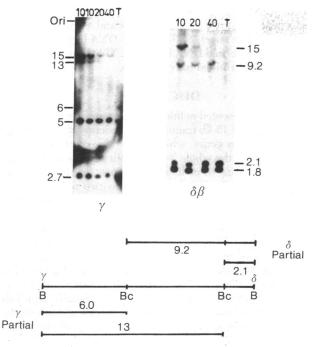


FIG. 3. Partial digests of the 15-kb BamHI fragment with Bcl I. BamHI-digested human DNA was digested for the indicated periods of time (10, 20, and 40 min and total digest of 3 hr, T) with Bcl I. Samples of 15 g were run on agarose gels, blotted, and hybridized. Y is hybridized with pHyG1 and $\delta\beta$ with the 5' probe derived from pH β G1. The map shows the origin of the various partial fragments; their sizes are tabulated in Table 1. Sizes are in kb.

In Fig. 4 we see that partial digests of the 15-kb BamHI fragment with Bgl II, hybridized with γ -gene probes, show not only the 6-kb total digest fragment, but also partial digest fragments of 8.0 and 10.5 kb. These fragments predict the same three Bgl II sites mapped from the δ and β partial digest fragments within the experimental error of the measurements.

Further evidence for the linkage of the δ and $^{\Lambda}\gamma$ genes was provided by analysis of human DNA digested with Kpn I. A single fragment of 46 kb was seen that hybridizes with both $\delta\beta$ -and γ -globin probes (Fig. 5A). One end of the Kpn I fragment has been mapped to be near the 3' end of the β -globin gene with double digests of Kpn I with Hpa I, BamHI, or EcoRI (Fig. 5B). The BamHI/Kpn I double digest showed a new fragment of 4.8 kb, whereas the 15-kb δ -globin and 1.8-kb β -globin fragments were uncleaved. This suggests that the Kpn I site is lo-

Table 1. Partial digests of the 15-kb BamHI fragment with Bcl I

Gene probed	Fragment size, kb			
	BamHI	Partial Bcl I/BamHI	Total Bcl I/BamHI	
γ	15.0	15.0		
		13.0		
		6.0	6.0	
	5.0	5.0	5.0	
	2.7	2.7	2.7	
$\delta \beta$, 5'	15.0	15.0		
regions		9.2		
		2.1	2.1	
	1.8	1.8	1.8	

Human DNA, digested as indicated, was hybridized with 32 P-labeled pH $_{\gamma}$ G1 or a 5' probe of pH $_{\beta}$ G1 DNA (see *Materials and Methods* and Fig. 3).

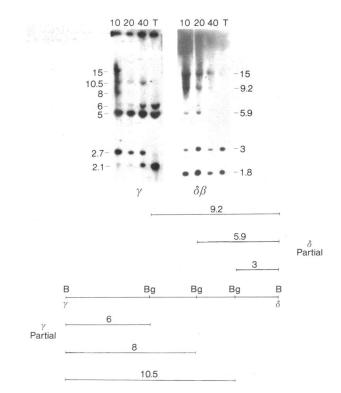


FIG. 4. Partial digests of the 15-kb BamHI fragment with Bgl II. BamHI-digested human DNA was digested with Bgl II for the times indicated (10, 20, and 40 min and a 3-hr total digest, T). The γ fragments were detected with pH γ G1 DNA; the β and δ fragments were detected with a 5′ pH β G1 probe. The map shows the origin of the partial bands; the fragment sizes are tabulated in Table 2. Sizes are in kb.

cated 4.8 kb to the 3' side of the intragenic BamHI site in the β -globin gene. This was confirmed by the fact that the 8.3-kb Hpa I β fragment was cleaved to give a 6.1-kb double-digest fragment and the 4-kb 3' EcoRI β -globin fragment was cut to give a 3.9-kb fragment. Kpn I did not cleave either the 10.3-kb Xba I fragment, which links the δ - and β -globin genes, or the 15-kb BamHI fragment or the EcoRI and BamHI fragments containing the γ -globin genes (data not shown).

The length required to contain all the DNA from the EcoRI site to the 5' side of the $^{G}\gamma$ -globin gene to the Kpn I site to the

Table 2. Partial digests of the 15-kb BamHI fragment with Bgl II

Gene probed	Fragment size, kb			
	BamHI	Partial Bgl II/BamHI	Total Bgl II/BamHI	
γ	15.0	15.0		
		10.5		
		8.0		
		6.0	6.0	
	5.0	5.0	5.0	
	2.7	2.7		
		2.1	2.1	
$\delta \beta$, 5'	15.0	15.0		
regions		9.2		
		5.9		
		3.0	3.0	
	1.8	1.8	1.8	

The human DNA was digested with the enzymes indicated as described in Fig. 4 and *Materials and Methods*. γ gene fragments were detected with ³²P-labeled pH γ G1 DNA and the 5' regions of the β and δ genes with a 5' pH β G1 probe.

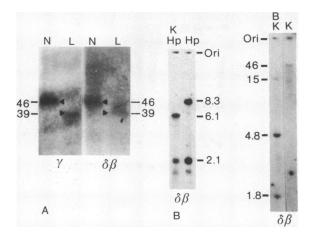


FIG. 5. Kpn I digestion of the human γ -, δ -, and β -globin genes. (A) Kpn I-cut normal DNA (N) or homozygous Hb^{Lepore} DNA (L) was digested with Kpn I and analyzed for γ -, δ -, and β -globin genes with $pH\gamma G1$ (γ) or $pH\beta G1$ ($\delta\beta$) probes. (B) Normal human DNA was cut with Hpa I (Hp), Hpa I plus Kpn I (K Hp), Kpn I (K), or BamHI plus Kpn I (K) and analyzed for δ - and β -globin genes with FK DNA digested with FK I or FK I was denatured before it was run on 1.2% agarose gels. Sizes are in K b.

3′ side of the β -globin gene (see Fig. 6) would be 37 kb if there were only a single linking 15-kb BamHI fragment, but it would be 52 kb if there were two such fragments linked in tandem. From the data presented above we know that the 46-kb Kpn I fragment must contain all of this DNA (and more), and we conclude that only a single 15-kb BamHI fragment is compatible with this fact.

It remains theoretically possible that there are two 46-kb Kpn I fragments. This possibility is eliminated by digests of DNA from a patient homozygous for Hb^{Lepore}. We have previously shown (6) that Hb^{Lepore} is caused by deletion of 6.6 kb of DNA, resulting in a fusion of the δ - and β -globin genes. We would, therefore, predict that the Kpn I globin gene fragment from Hb^{Lepore} DNA would be smaller by 6.6 kb than the normal fragment, as is demonstrated in Fig. 5A. A single globin gene fragment of 39 kb was seen. The length of DNA required to contain all the globin gene fragments from a Hb^{Lepore} patient would be 30.4 kb (37 - 6.6 kb) if there were a single 15-kb BamHI fragment and 45.4 kb (52 - 6.6 kb) if there were two tandem 15-kb BamHI fragments. The size markers used allowed accurate size estimation in this region and showed that there was a single 15-kb BamHI fragment.

In addition, BamHI/Kpn I digests of Hb^{Lepore} DNA contained a normal 4.8-kb β double digest fragment (data not presented), indicating that the change in mobility of the Kpn I globin gene fragment could only be caused by the 6.6-kb deletion.

These experiments prove that there is only a single Kpn I globin gene fragment because both β - and γ -specific probes hybridized to the novel 39-kb Hb^{Lepore} DNA fragment as well as to the 46-kb normal DNA fragment. This allows the map in Fig. 6 to be determined unequivocally.

DISCUSSION

The evidence presented in this paper demonstrates that BamHI generates a single 15-kb fragment that contains part of both the $^{
m A}\gamma$ - and δ -globin genes, which allows the construction of a physical map of the whole β -like globin gene locus. We have shown this by using three lines of evidence. (i) Attempts to resolve the 15-kb BamHI fragment into two separate 15-kb fragments by using RPC-5 chromatography were not successful. All other γ -, δ -, and β -globin gene fragments were resolved by this process. (ii) The pattern of Hpa I, Bgl II, and Bcl I sites predicted to lie within the 15-kb BamHI fragment is identical when the prediction is made from the γ - or from the δ -globin gene maps. We believe that it is highly unlikely that two separate 15-kb fragments generated by BamHI would have identical sites for three other enzymes. (iii) Finally, Kpn I digests of normal and HbLepore DNA generate a single, large, globin gene fragment that contains the γ - and the δ - and β -globin genes. The size of this fragment does not allow the existence of two 15-kb fragments spanning the $^{A}\gamma$ - δ gene gap.

The main features of the β -like globin locus are shown in Fig. 6. The two γ -globin genes are separated by 3.5 kb, the $^{\rm A}\gamma$ - and δ -globin genes by 13.5 kb, and the δ - and β -globin genes by 5.5 kb. All four genes contain intervening sequences of 800–1000 base pairs and all four genes are transcribed from the same DNA strand. The physical map of the four genes described here raises a number of questions. Is it significant that all four genes are transcribed in the same relative direction? Is it significant that the two pairs of genes whose expression is linked ($^{\rm A}\gamma$ and $^{\rm G}\gamma$; δ and β) are physically closer than those whose expression is not? The isolation of this entire region in a recombinant should make it possible to correlate structure with function more directly

The determination of the physical distance between the β and γ -globin genes allows direct correlation of recombination
rates with DNA length and also places size constraints upon
models of evolution by gene duplication.

The same ${}^{\Lambda}\gamma$ - δ intergene distance has recently been determined independently by T. Maniatis and his colleagues using a different approach (10).

We thank Mieke Lupker-Wille for a gift of *Hpa* I. We thank Caroline Rogers for preparing the manuscript. This work was supported by grants to R.A.F. from The Netherlands Foundation for Chemical Research (SON) with financial aid from The Netherlands

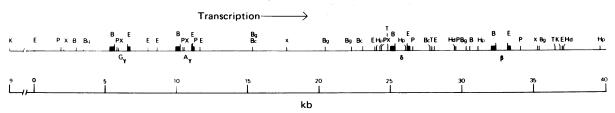


FIG. 6. A map of the linked human adult and fetal globin gene loci. Additional data are taken from refs. 5 and 6 and unpublished data. Not all sites for a given enzyme are shown, and only restriction enzyme sites that generate fragments that contain globin genes are represented here. Coding regions are shown as filled boxes. The possibility that these coding regions are split by further intervening sequences cannot be excluded from the published data (5, 6). The distances in the δ - and β -globin gene map differ slightly from those published in ref. 6. This is the result of recalibration of the larger (>3 kb) globin gene fragments. This results in a δ - β intergene distance of 5.5 kb rather than 7 kb, as originally published. Full details of these fragment sizes are available from R.A.F. on request. B, BamHI; Bc, Bcl I; Bg, Bgl II; E, EcoRI; P, Pst I; X, Xba I; Hp, Hpa I; and K, Kpn I.

Organization for the Advancement of Pure Research (ZWO) and to R.W. from the U.K. Medical Research Council and from the U.S. National Institutes of Health (1R01AM20125-01A1). A travel grant from the North Atlantic Treaty Organization made the collaboration possible.

- Weatherall, D. J., Pembrey, M. E. & Prithcard, J. (1974) Clin. Haematol. 3, 467-508.
- 2. Lehmann, H. & Huntsman, R. G. (1974) Man's Haemoglobins Including the Haemoglobinopathies and Their Investigation (North-Holland, Amsterdam), 2nd Ed.
- Schroeder, W. A., Huisman, T. H. J., Shelton, R., Shelton, J. B., Kleihauer, E. F., Dozy, A. M. & Robberson, B. (1968) Proc. Natl. Acad. Sci. USA 75, 882–886.

- 4. Wood, W. G. (1976) Br. Med. Bull. 32, 282-287.
- Little, P. F. R., Flavell, R. A., Kooter, J. M., Annison, G. & Williamson, R. (1979) Nature (London) 278, 227–231.
- Flavell, R. A., Kooter, J. M., De Boer, E., Little, P. F. R. & Williamson, R. (1978) Cell 15, 25–41.
- 7. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
- 8. Jeffreys, A. J. & Flavell, R. A. (1977) Cell 12, 1097-1108.
- Little, P. F. R., Curtis, P., Coutelle, C., Van Den Berg, J., Dalgleish, R., Malcolm, S., Courtney, M., Westaway, D. & Williamson, R. (1978) Nature (London) 273, 640–643.
- Fritsch, E. F., Lawn, R. M. & Maniatis, T. (1979) Nature (London) 279, 598-603.