Volume 8 Number 19 1980 Nucleic Acids Research

The 5' flanking region of human ε-globin gene

Francisco E.Baralle[1*], Carol C.Shoulders[1], Steven Goodbourn[2], Alec Jeffreys[3] and Nicholas J.Proudfoot[4]

[1]MRC Laboratory of Molecular Biology, Hills Road, Cambridge, [2]MRC Unit for Molecular Haematology, John Radcliffe Hospital, Oxford, [3]Department of Genetics, University of Leicester, UK, and [4]Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA

Received 25 July 1980

ABSTRACT

The structural analysis of the 2.0 kb region upstream from the ε-globin gene has been carried out. A genomic DNA map around the gene was worked out in some detail to ensure that the cloned DNA was representative of the actual chromosomal arrangement. Furthermore, a new technique was developed to precisely map a reiterated DNA sequence present 1.5 kb to the 5' side of the gene. The complete nucleotide sequence of the 2.0 kb 5' flanking region was then determined and overlapped with the gene. The sequence included the reiterated DNA sequence which is homologous to the so-called AluI family of repeats. Unusual stretches of sequence 50 nucleotides long, where A + T represent about 90% of the bases, are present at both the 5' and 3' sides of the repeat.

INTRODUCTION

The human globin gene locus has recently been the subject of extensive structural studies. In particular, the nucleotide sequence of the ε, $^{G}\gamma$, $^{A}\gamma$, δ and β globin structural genes has been completed [1-5] and a good knowledge of the restriction enzyme sites in the extragenic regions is available [6-8]. However, more detailed information about the structure of the flanking regions seems desirable as there is a growing body of evidence that they influence the regulation of gene expression [7-11]. A prominent feature of the extragenic regions in most eucaryotic genomes, including human, is that 300 nucleotides long repeats [12-15] are interspersed with longer single copy sequences. Reiterated DNA sequences were indeed found in the flanking region of ε, δ, γ and β human globin genes [8,15]. The repeats 5' to the δ gene and 3' to the β gene are located within regions apparently involved in the cis acting suppression of foetal gene expression in adults [9-11]. Furthermore, it has been shown that some of the repeated sequences are transcribed "in vitro" by RNA polymerase III producing discrete RNA molecules [15].

We wish to describe an extension of our previous ε-globin gene studies [1,16,17] to the 5' flanking region. Initially a genomic DNA map was worked

out in some detail to ensure that the cloned 8 kb fragment was representative of the actual chromosomal arrangement. Then a study of the organisation of the repeated sequences in the clone was carried out. Finally, the complete nucleotide sequence of the 2.0 kb 5' flanking region was determined. It included the reiterated DNA sequence and overlapped with the 5' end of the ε-globin gene, whose complete sequence has been previously reported [1].

MATERIALS AND METHODS

(a) Blotting experiments

Restriction enzyme digests, blotting and hybridisation were as previously described [18] except that hybridisations were performed at 10 ng of ^{32}P probe per ml of hybridisation solution supplemented with 9% Dextran sulphate (Sigma, M.W. 500,000). The molecular weight of fragments were determined using λ HindIII and pBR322 fragments as markers [18]. In the experiments shown in Table 1 and Figures 1 and 2 the DNA fragments were denatured in alkali before the agarose gel electrophoresis, whereas in the experiments shown in Figures 3 and 4 the DNA fragments were denatured after the electrophoresis. This should be noted, because the estimated size of the fragments is consistently lower using the latter method. For the genomic mapping three different probes were used (see Results). The probe used for mapping the reiterated DNA sequences was nick-translated total genomic DNA.

(b) Fine mapping of reiterated DNA sequences

A new fast mapping technique that can be applied to restriction enzyme fragments which have 5' overhangs was developed. The methodology followed to map the HinfI fragments of the EcoRI/BamHI 2.0 kb fragment will be described as an example. An EcoRI/BamHI digest of the cloned 8 kb HindIII fragment [16,17] was fractionated in a 6% thin layer acrylamide gel [19] and the 2.0 kb EcoRI/BamHI band was eluted [20]. The fragment was then treated with HinfI and the digestion products labelled by filling in the 5' overhangs using E. coli DNA polymerase (Klenow fragment) and α-^{32}P-dATP as the only source of nucleotide triphosphates. The reaction mixture was fractionated in an 8% acrylamide gel and the bands eluted, ethanol precipitated and redissolved in 20 µl of 20 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 100 mM NaCl and 2 mM DTT. A 10-fold molar excess of single stranded DNA from the subclone M13ε3.7+ [1] was added, the mixture was heated for 5 min at 100°C in a sealed capillary, and then annealed for 60 min at 67°C. The contents of the capillary were blown into 20 µl of distilled water, the four deoxynucleoside triphosphates to a final concentration of 500 µM and 2 µl (1 unit) of Klenow

DNA polymerase were added. The reaction mixture was incubated at room temperature for 30 min and the enzyme was inactivated by heating at 70°C for 10 min. After cooling, HinfI was added and the mixture incubated for 60 min at 37°C; the digest was subsequently fractionated on an 8% acrylamide gel (see Fig. 4). Once the relative position of the HinfI fragments in the region was established the reiterated DNA sequence was mapped by hybridising nick-translated total human genomic DNA to the individual HinfI fragments immobilised on diazobenzyloxymethyl (DBM) paper [21].

(c) Sequencing methodology

The EcoRI/BamHI 2.0 kb fragment was either isolated from λε [1,18] or from the subclone pH 1.8 (generously provided by E. Fritsch, see ref. 8). Most of the sequence was determined by the chemical degradation procedure [20]. The EcoRI, BglII, PvuII and the HinfI sites were used for labelling either at the 3' end by "filling in" or at the 5' end with polynucleotide kinase [20]. The HinfI fragments were strand separated by annealing them to the M13ε3.7 subclones as previously described [1]. The region between the EcoRI and BglII sites was cross checked by labelling partial HaeIII digests of single strand DNA from M13ε3.7+ as previously described [1]. The 3' end of the HinfI fragment number 4 was also cross checked by the dideoxy technique [22] using M13ε3.7+ as template [23] and the HinfI fragment number 10 as a primer.

RESULTS AND DISCUSSION

(a) Mapping of human genomic DNA around the ε-globin gene

The restriction enzyme analysis of the human chromosomal DNA around the ε-globin gene was carried out using three different probes. A gene probe consisted of the BamHI 700 bp long fragment that contained the 5' end of the ε-globin gene. The 3' BamHI site is 20 nucleotides to the 5' end of the middle exon - large intron junction [1]. The 5' extragenic probe was a 500 bp HindIII/XbaI fragment situated about 5 kb to the 5' end of the gene [16] and the 3' extragenic probe was an EcoRI/HindIII fragment situated about 200 bp to the 3' end of the poly(A) addition site [1,16]. The results of the hybridisation experiments are summarised in Table 1 and some examples of the blots are shown in Figure 1. As can be seen, the ε probe hybridises to the 5' region of β, $^{G}\gamma$, $^{A}\gamma$, δ and ε at low stringencies (0.5 x SSC, 65°C), at higher stringencies all except the specific ε bands are preferentially lost (see Fig. 1). There are additional bands seen in the BclI, PstI and BamHI digests that are not assignable to γ, δ, β or ε. As they were also faintly

Table 1. Sizes (in kb) of DNA fragments containing the ε-globin gene or neighbouring sequences in restriction endonuclease digests of total human DNA.

Restriction endonuclease	Probe		
	ε-gene	5' extragenic	3' extragenic
BamHI	15.5 5.5 2.7 2.0 0.7[a]	15.9	15.2
BclI	19.0 12.1 11.5[a] 8.6[b] 4.9	11.5	11.5
EcoRI	7.6 6.5 4.5[a] 2.9 2.6	7.9	5.5
KpnI	15.5[a]	15.5	15.5 8.0
PstI	15.3[a] 5.3 4.5 2.6	15.3	15.3
BamHI + BclI	8.6[b] 5.5 2.7 2.0 0.7[a]	7.6	3.1
BamHI + EcoRI	7.9[b] 2.7 2.0 1.9 1.6 0.7[a]	7.9	5.5
BamHI + KpnI	15.5 5.5 2.7 2.0 0.7[a]	13.0	8.0 1.6
BamHI + PstI	5.5 4.4[b] 2.7 2.0 1.0 0.7[a]	10.7	4.7
BclI + KpnI	19.0 12.1 10.0[a] 8.6[b] 4.9	10.0	10.0 1.2
BclI + PstI	11.5[a] 5.3 4.5 2.6	11.5	11.5
EcoRI + KpnI	7.6 6.5 4.5[a] 2.9 2.6	7.9	5.3
EcoRI + PstI	4.5[a] 2.1	7.9	3.3
KpnI + PstI	12.9[a] 5.3 4.5 2.6	12.9	12.9 2.9

10 μg samples of human placental DNA were cleaved with the indicated combinations of restriction endonucleases, denatured with alkali, electrophoresed through a 0.8% agarose gel and transferred to a nitrocellulose filter. The filter was hybridised with ^{32}P-labelled 700 bp cloned BamHI fragments containing the 5' end of the ε-globin gene plus additional 5' extragenic sequences (see Fig. 2). The filter was given a post-hybridization wash in 0.5 x SSC at 65° and labelled bands were detected by autoradiography. Novel double digest fragments, seen in neither single digest are underlined. The filter was then washed at increased stringency (0.1 x SSC at 65°) and remaining labelled fragments (a) containing the ε-globin gene were again detected by autoradiography. At 0.5 x SSC and 65°, additional labelled components could be detected; in all cases except those marked (b) these fragments can be assigned to the $^{G}\gamma$, $^{A}\gamma$, δ or β-globin genes [25,26].

A second identical filter was hybridised with ^{32}P-labelled 500 bp cloned HindIII + XbaI double digest fragment containing sequences 5 kb on the 5' side of the ε-globin gene. The filter was given a post-hybridisation wash in 0.3 x SSC at 65° and remaining labelled bands detected by autoradiography. Label was then removed by washing for 1 hr in water at 65°, and the filter rehybridised with ^{32}P-labelled 700 bp cloned HindIII + EcoRI fragment containing extragenic sequences 0.4-1.2 kb from the 3' end of the ε-globin gene. Labelled fragments were detected as before.

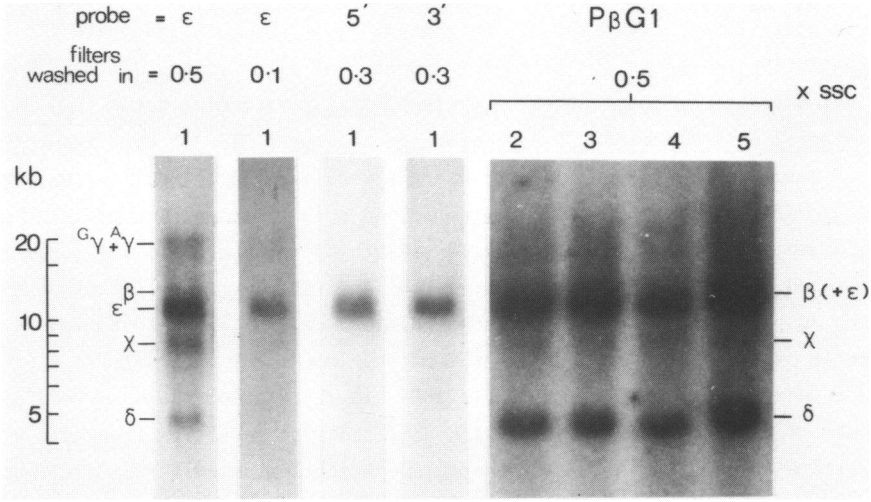


Figure 1. DNA fragments containing the ε-globin gene or sequences neighbouring this gene in human DNA digested with restriction endonuclease BclI. DNA isolated from human placenta (lane 1) or from blood obtained from four unrelated individuals (lanes 2-5) was cleaved with restriction endonuclease BclI. 10 μg samples were denatured with alkali, electrophoresed through an 0.8% agarose gel and transferred to a nitrocellulose filter. Filters were hybridised with the indicated ^{32}P-labelled DNAs in 1 x SSC at 65° (see Methods). After hybridisation, label was removed by washing in 1 x SSC at 65° and filters were given a further high stringency wash in the indicated concentration of SSC. The origin of the human probes used is shown in Figure 2. Plasmid pβG1 contains a cDNA copy of rabbit β-globin mRNA [24]. The positions of fragments containing the β, δ, $^{G}\gamma$, $^{A}\gamma$ and ε-globin genes are shown on the autoradiograph. Restriction endonuclease BclI gives a single human DNA fragment containing both the $^{G}\gamma$ and $^{A}\gamma$ globin genes. In addition there is a 8.6 kb fragment (marked X) which is not assignable to the known β-related globin genes.

detected in the BclI digest of four individuals using pβG1 (a rabbit β-globin cDNA plasmid, ref. 24) as a probe the possibility of these bands being polymorphic variants can be safely ruled out. Since the bands are detected by pβG1 they must represent an additional β-related globin gene. These results are consistent with our previous observation of non-assignable bands hybridising to pβG1 in HindIII and EcoRI digests of human DNA [16]. Recently, E. Fritsch et al. [8] identified these bands as belonging to a new family of sequences, the pseudo β-globin genes. The 5' and 3' extragenic probes (see Table 1 and Fig. 1) seem to act as specific single copy sequences and hybridise to single bands in the digests tried. Hence the DNA in these spacer or extragenic regions is unique and not repetitive. All the digests

and double digests shown in Table 1 and Figure 1 can be used to construct a unique restriction enzyme map of the region. Such a map is presented schematically in Figure 2. It covers a region of about 35 kb. It confirms and extends the map constructed by Fritsch et al. using cloned DNA [8].

(b) Mapping of reiterated DNA sequences on the cloned HindIII 8kb fragment

A series of restriction enzyme digests were carried out on the isolated 8 kb HindIII fragment. The bands were blotted onto a nitrocellulose filter and hybridised to a probe prepared by nick-translated total genomic DNA (only the reiterated DNA sequences will be present at sufficiently high concentrations to act as probe at the hybridisation conditions used). Figure 3 shows examples of these blots; it is clear from this data that at least two repeats are present in the clone. The EcoRI/BamHI 2.0 kb fragment was selected for further studies. It was digested with HinfI and the resulting fragments were labelled at the 3' end and fractionated in an acrylamide gel. The fragments were eluted, extended with Klenow DNA polymerase using M13ε3.7+ as template and digested again with HinfI. These operations transfer the label to the nearest neighbour fragment (see following diagram).

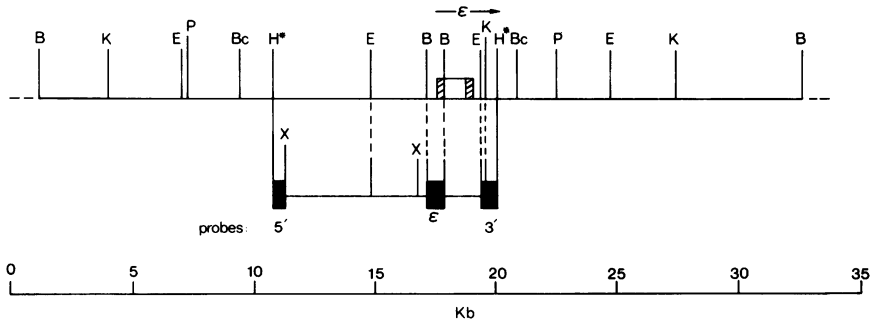
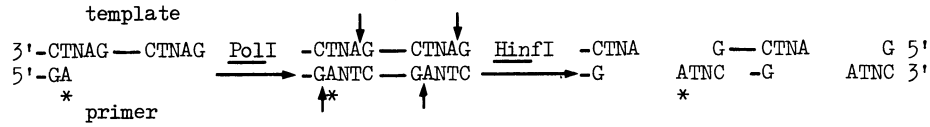


Figure 2. A physical map of restriction endonuclease cleavage sites around the ε-globin gene. This map was constructed from the mapping data in Table 1 and shows the ε-globin gene (hatched boxes) and cleavage sites for restriction endonucleases BamHI (B), BclI (Bc), EcoRI (E), KpnI (K) and PstI (P). This map is aligned with the cloned HindIII (H) fragment and with cleavage sites within this fragment. The positions of sites marked * were not established by analysis of human DNA digests and are taken from positions on the cloned ε-globin DNA fragment. Fragments isolated from this cloned segment are shown as solid boxes and were used as hybridisation probes to establish the physical map. Human DNA digested with XbaI (X) was not analysed.



The results obtained are shown in Figure 4, lanes 1 to 10. Conventionally the fragments were ordered in a 5'—3' orientation being the 5' end farthest away from the ε-globin gene. It was known that fragment 2 was the HinfI/BamHI fragment and hence the closest to the ε-globin gene. In lane 2 it is clearly visible that the original fragment has transferred its label to fragment 6. According to our convention 6 is then positioned immediately to the 5' end of 2. Routinely only a fainter band was present at the position of the fragment used as primer (in this case fragment 2); equally intense bands should be expected because the complementary strand cannot be extended. We assume that the label is lost by exchange with cold dATP during the extension reaction. The faint larger bands visible in some of the lanes may represent partial HinfI cleavages of the extended fragments. Fragment 7 is the EcoRI/

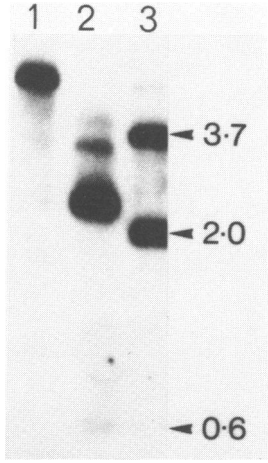


Figure 3. The HindIII 8 kb cloned fragment (see Fig. 2 and refs. 1,16) was digested with the following restriction enzymes: lane 1, BamHI; lane 2, BglII; lane 3, EcoRI/BamHI, and fractionated in an 0.8% agarose gel. The fragments were visualised by staining with ethidium bromide (not shown), blotted and hybridised to ^{32}P-labelled total genomic DNA (see Materials and Methods). Hybridisation of the probe is observed to the following fragments: EcoRI/BamHI, 3.7 and 2.0 kb; BamHI/HindIII, 5.5 kb; BglII, 2.2 kb and 0.6 kb (weakly); BglII/HindIII, 3.5 kb. No hybridisation is observed to the BamHI 0.7 kb, BamHI/EcoRI 1.3 kb, EcoRI/HindIII 0.7 kb, BglII/HindIII 1.9 kb. This data shows unequivocally that at least two repeats are present in the HindIII 8 kb fragment.

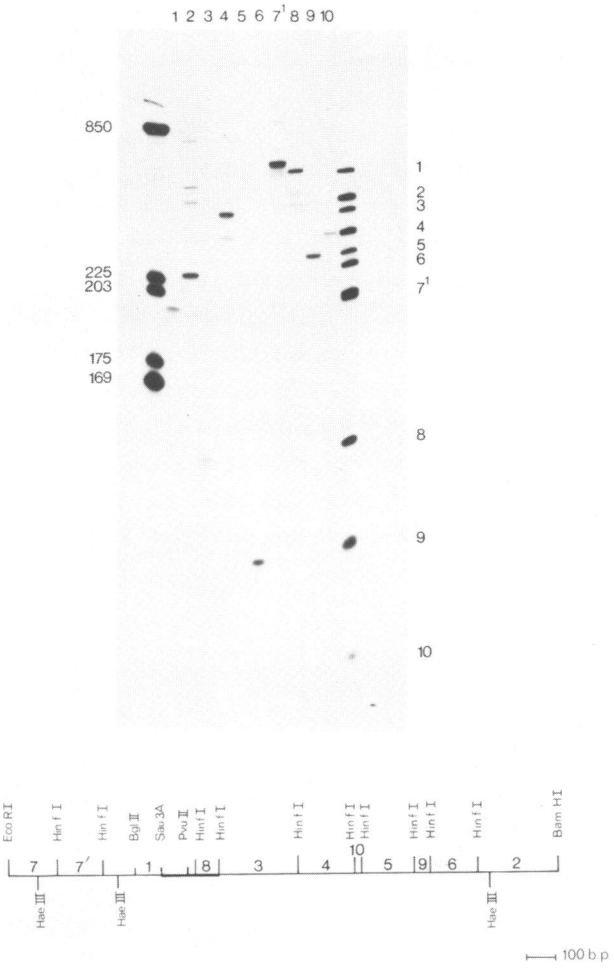


Figure 4. Nearest neighbour fragment mapping of the EcoRI/BamHI 2.0 kb fragment. The HinfI fragments of the EcoRI/BamHI 2.0 kb (far right lane of the gel) are numbered 1-10. 7^1 means that two fragments 7 and 7' are present at that position. The order of the fragments can be directly deduced from the gel, as described in the text. The restriction enzyme map shown was constructed by combining the nearest neighbour fragment data deduced from the gel with more classical mapping of other restriction enzyme sites (data not shown). The thicker black line denotes the approximate position of the repeated DNA sequences.

HinfI fragment so its extension goes beyond the 2.0 kb and a large band is visible, possibly corresponding to the first HinfI site in the M13 DNA. As 7 and 7' were not separated before annealing to M13mp3.7+ it is hard to explain why there is not an equally intense band corresponding to the fragment 7 produced by the extension of 7'. In fact, the conclusive evidence for ordering these fragments came only after the complete sequence of the region was determined. The HinfI fragments 1 and 8 were identified as the ones containing the reiterated DNA sequence by hybridisation of the individual HinfI fragments immobilised on DBM paper to nick-translated genomic DNA (data not shown). Lane 5 showed a faint band with the mobility of fragment 10.

(c) The complete nucleotide sequence of the EcoRI/BamHI 2.0 kb fragment

Figure 5 shows the primary structure of the 5' flanking region of human ε-globin gene. The sequence at positions 1-750 and 1650-1850 was determined in two clones: λε [16] and pH 1.8 (constructed by E. Fritsch). As they were isolated from different individuals we had the opportunity to analyse polymorphism in the region at the nucleotide level. No single base changes were detected; however, the dinucleotide C-A at position 384-385 was deleted in pH 1.8. It is not possible to decide if this difference is a real polymorphism or an artifact produced during the cloning or propagation of the clones. In view of independent evidence on the rate of polymorphism in the human genome [18] the former alternative seems more likely.

The region between nucleotides 450-750 present homologies with the AluI family of repeated DNA sequences [28]. The start and end of the repeat cannot be precisely defined until the homologous regions in other parts of the genome are sequenced. The characteristic AluI restriction site is at position 593, which is 1485 bp to the 5' side of the ε-globin A-U-G initiator codon. The AluI family of repeats presents homologies with most of the prominent oligonucleotides found in the repetitive double strand region of HeLa cell hnRNA [28-32]. Furthermore, sequence similarities have been noted between selected portions of the AluI family and several other RNA or DNA sequences which are known or suspected to be involved in DNA replication, transcription control and mRNA processing [32]. In this respect it is interesting to note that the repeat is flanked by A-T rich regions. In fact, A + T accounts for 92% of the bases between position 200-250 including a run of 21 residues (see Fig. 6) and between position 707-758 A + T account for 88% of the residues.

Our mapping data showed that at least another repeat is present to the left of the central EcoRI site of the HindIII 8.0 kb fragment (see Figs. 2 and 3). Recent electron microscopy studies have shown that self annealed molecules of the HindIII 8.0 kb fragment cloned in either λ788 or pAT153 form a loop of single stranded DNA and a duplex DNA stem with an estimated size of 227 ± 39 base pairs [33]. This foldback structure revealed the presence of a pair of inverted repeat sequences which was able to hybridise intramolecularly. It can thus be inferred that most of the nucleotides between position 450 and 750 (Fig. 5) form one of the strands of the stem.

In vitro transcription studies have shown that the RNA synthesised from these regions by RNA polymerase III contains segments of unique nucleotide sequence in addition to the repeat [15; E. Fritsch, personal communication;

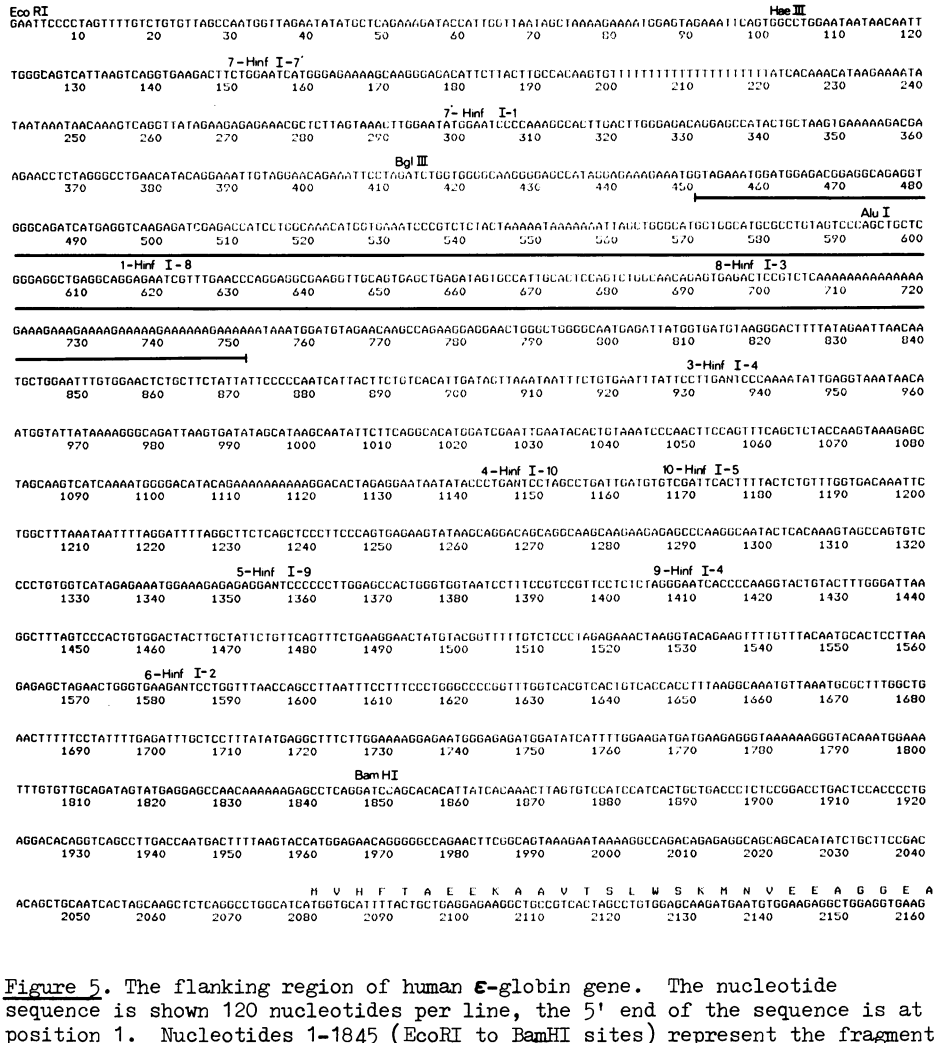


Figure 5. The flanking region of human ε-globin gene. The nucleotide sequence is shown 120 nucleotides per line, the 5' end of the sequence is at position 1. Nucleotides 1-1845 (EcoRI to BamHI sites) represent the fragment whose restriction enzyme map is shown in Figure 4. Nucleotides 1845-2160 belong to the BamHI 0.7 kb fragment previously sequenced (see Fig. 2 and ref. 1). The first 28 amino acids of ε-globin are indicated above each coding sequence using the amino acid abbreviation adopted by Dayhoff [27]. The more relevant restriction sites are indicated. The numbers at each side of the HinfI sites correspond to the fragments described in Figure 4. The sequence underlined (positions 450-750) contains the repeat. It should be noted that the ends of the reiterated DNA sequence are not well defined (see text and ref. 28). The HinfI sites at positions 923, 1134, 1169, 1353, 1398 and 1581 were not read through when sequencing but they are unequivocally overlapped by the nearest neighbour fragment mapping (see Fig. 4). N denotes an unknown nucleotide. The sequences between positions 930 and 980, 1145 and 1175, 1350 and 1415, 1580 and 1590, 1620 and 1650 were only read in one strand and therefore should be considered tentative (about 98% certain).

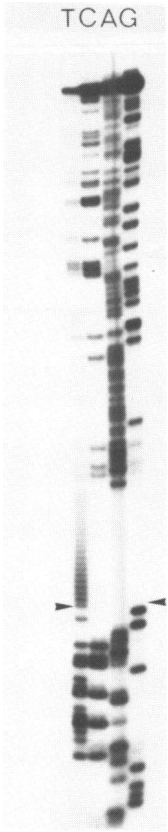


Figure 6. Autoradiograph of a gel analysing the products yielded by the four base-specific chemical degradations [20]. T, C, A and G denote T + C, C-specific, A > C, and G-specific cleavages. The arrows correspond to position 200 of the sequence described in Figure 5 and the unusual 92% A + T content of this region is clearly evident.

Tocchini Valentini and Baralle, unpublished results). Therefore it was possible that these unique sequences could be somehow related to the gene downstream. However, a comparison of the nucleotide sequence of the 5' flanking region with the ε-globin gene has failed to show any significant primary structure relationship.

The knowledge of the complete nucleotide sequence of the 5' flanking region of the ε-globin gene allowed us to devise several functional experiments now in progress aimed to elucidate the biological significance of the region.

*Present address: Sir William Dunn School of Pathology, South Parks Road, Oxford OX1 3RE, UK

REFERENCES

1. Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980). Cell 21, 621-626.

2. Lawn, R.M., Efstratiadis, A., O'Connell, C. and Maniatis, T. (1980). Cell 21, 647-651.
3. Spritz, R.A., DeRiel, J.K., Forget, B.G. and Weissman, S.M. (1980). Cell 21, 638-646.
4. Slightom, J.L., Blechl, A.E. and Smithies, O. (1980). Cell 21, 627-637.
5. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980). Cell 21, 653-668.
6. Bernards, R., Little, P.F.R., Annison, G., Williamson, R. and Flavell, R.A. (1979). Proc. Nat. Acad. Sci. USA 76, 4827-4831.
7. Tuan, D., Biro, P.A., DeRiel, J.K., Lagorus, H. and Forget, B.G. (1979). Nucleic Acids Res. 6, 2519-2544.
8. Fritsch, E.F., Lawn, R.M. and Maniatis, T. (1980). Cell 19, 959-972.
9. Weatherall, D.J. and Clegg, J.B. (1979). Cell 16, 467-479.
10. Fritsch, E.F., Lawn, R.M. and Maniatis, T. (1979). Nature 279, 598-603.
11. Bernards, R., Kooter, J.M. and Flavell, R.A. (1979). Gene 6, 265-280.
12. Davidson, E.H. and Britten, R.J. (1979). Science 204, 1052-1059.
13. Schmid, C.W. and Deininger, P.L. (1975). Cell 6, 345-358.
14. Deininger, P.L. and Schmid, C.W. (1976). J. Mol. Biol. 106, 773-790.
15. Duncan, C., Biro, P.A., Chowdary, P.V., Elder, J.T., Wong, R.R.C., Forget, B.G., DeRiel, J.K. and Weissman, S.M. (1979). Proc. Nat. Acad. Sci. USA 76, 5095-5099.
16. Proudfoot, N.J. and Baralle, F.E. (1979). Proc. Nat. Acad. Sci. USA 76, 5435-5439.
17. Baralle, F.E., Proudfoot, N.J. and Clegg, J.B. (1980). Ann. N.Y. Acad. Sci., in press.
18. Jeffreys, A.J. (1979). Cell 18, 1-10.
19. Sanger, F. and Coulson, A.R. (1978). FEBS Lett. 87, 107-110.
20. Maxam, A.M. and Gilbert, W. (1977). Proc. Nat. Acad. Sci. USA 74, 560-564.
21. Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977). Proc. Nat. Acad. Sci. USA 74, 5350-5354.
22. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). Proc. Nat. Acad. Sci. USA 74, 5463-5467.
23. Schreier, P.H. and Cortese, R. (1979). J. Mol. Biol. 129, 169-172.
24. Maniatis, T., Kee, S.G., Efstratiadis, A. and Kafatos, F.C. (1976). Cell 8, 163-182.
25. Flavell, R.A., Kooter, J.M., DeBoer, E., Little, P.F.R. and Williamson, R. (1978). Cell 15, 25-41.
26. Little, P.F.R., Flavell, R.A., Kooter, J.M., Annison, G. and Williamson, R. (1979). Nature 278, 227-231.
27. Dayhoff, M.O. (1969). Atlas of protein sequence and structure. National Biomedical Research Foundation, Silver Spring, Maryland, USA.
28. Rubin, C.M., Houck, C.M., Deininger, P.L., Friedmann, T. and Schmid, C.W. (1980). Nature 284, 372-374.
29. Robertson, H.D., Dickson, E. and Jelinek, W.R. (1977). J. Mol. Biol. 115, 571-589.
30. Jelinek, W.R. (1977). J. Mol. Biol. 115, 591-600.
31. Jelinek, W.R., Evans, R., Wilson, M., Salditt-Georgieff, M. and Darnell, J.E. (1978). Biochemistry 17, 2776-2783.
32. Jelinek, W.R., Toomey, T.P., Leinwand, L., Duncan, C.H., Biro, P.A., Chowdary, P.V., Weissman, S.M., Rubin, C.M., Houck, C.M. Deininger, P.L. and Schmid, C.W. (1980). Proc. Nat. Acad. Sci. USA 1398-1402.
33. Coggins, L.W., Grindley, G.J., Voss, J.K., Slater, A.-A., Montagu, P., Stinson, M.A. and Paul, J. (1980). Nucleic Acids Res., in press.