Volume 12 Number 23 1984 Nucleic Acids Research

The chromatin structure of the human ε globin gene: nuclease hypersensitive sites correlate with multiple initiation sites of transcription

Jing-de Zhu*, Maggi Allan+ and John Paul+

Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, UK

Received 24 July 1984; Revised and Accepted 30 October 1984

ABSTRACT

We have mapped sites in chromatin flanking the ε-globin gene in the K562 cell which are hypersensitive to digestion with DNAseI, micrococcal nuclease and S1 nuclease. Many of those in the 5' flanking region correspond to minor upstream transcriptional starts. However, one prominent site occurs upstream of the boundary of transcription; it maps to a region with an unusual DNA sequence. In baby hamster kidney cells stably transformed with recombinant DNA containing the human ε-globin gene and in Cos 7 cells transiently transfected with DNA containing the ε-globin gene, hypersensitive sites can be demonstrated.

INTRODUCTION

Active genes are distinguished from inactive genes in chromatin both by an elevated overall sensitivity to DNAase I and by discrete localized regions, commonly at or near the 5' ends, which are much more sensitive to various nucleases and are more accessible to DNA binding compounds. These are referred to as "hypersensitive sites" (HS) (1-3). They have been identified in the 5' flanking region of genes encoding chick β-globin (4-6), Drosophila heat shock proteins (7) and histones (8,9), rat prepro-insulin II (10), HSV thymidine kinase (11), chicken conalbumin (12), a Drosophila glue protein (13), the major chicken vitellogenin gene (14) and on SV40 and polyoma minichromosomes (15,16). They frequently occur within a region which extends several hundred base pairs 5' to the cap site (5,7,9,10,11) and the development of HS around genes correlates with their transcription (4,10,13,14). It has been shown by the use of mutants that sequences in the region of HS are necessary for the expression of the Drosophila glue protein gene (17).

These features suggest that HS could be direct manifestations of interaction of regulatory proteins or RNA polymerases with promoter regions. This idea gains support from the recent demonstration by Emerson and Felsenfeld (18) that a protein from erythroid cell nuclei can bind

specifically to the HS 5' to the chick β-globin gene. However, HS also occur at sites in active chromatin where no promoter activity has been shown. Hence, their precise nature and function is still a matter for speculation.

Recently Allan et al (19,20) have shown that transcription of the human ε-globin gene in both the normal human embryo and the K562 cell, initiates at multiple upstream sites as well as from the canonical cap site. At least some of the resulting transcripts are capped and the generality of this observation has been established by the identification of alternative transcripts which initiate upstream from human β- and γ-globin genes (21-23).

In a previous publication (20) we mentioned that we had found that upstream starts in the human ε-globin gene coincide with regions of DNAseI hypersensitivity and in this paper we present the data and related findings. We used the indirect end-labelling technique (7) to map the HS around the human ε-globin gene in K562 nuclei and describe their occurrence at or near initiation sites of ε-globin transcripts both at the major cap site and also at positions about 200 bp, 850 bp, 1.45 kb and 4.5 kb upstream from it which appear to coincide with minor upstream origins of transcription. Another HS is found 6.3 kb upstream and this occurs at or very near to a long sequence of alternating purines and pyrimidines.

Experimental procedures

Cells and nuclei

Log phase K562 cells were grown in stirred suspension cultures to 10^5 cells/ml in a modified Ham's F12 medium (Flow Laboratories) supplemented with 15% horse serum (Gibco Europe Ltd). They were induced by growing in the same medium containing 0.1 mM hemin for 5 days. EJ and HES cell lines were grown in the same medium with 15% fetal calf serum. The cell line B202 was derived by Dr D Spandidos from a thymidine kinase deficient BHK cell which was stably transformed by treatment with a fusion plasmid containing herpes simplex virus thymidine kinase and the ε-globin gene; it was grown up in the same medium supplemented with HAT (24). Nuclei from K562 cells were prepared according to Pays and Gilmour (25). Preparation of nuclei from Cos 7 and B202 cell lines was carried out according to the procedure of Groudine and Weintraub (27).

Transient expression

Cos 7 cells were grown on Ham's F12 medium lacking hypoxanthine and thymidine and supplemented with essential amino acids and 10% fetal calf serum. Twenty-four hours prior to transfection the cultures were split to give 50% confluent cultures and medium was replaced 3-4 hours before addition

of DNA. A calcium phosphate precipitate was formed as described by Wigler et al (26) and cells were exposed to the precipitate for 14-16 hours after which the medium was replaced. Cells were trypsinized after incubation for a further 32 hours.

Nuclease digestion

Nuclei were either used directly for nuclease digestion studies or stored at -70° in a storage buffer (50mM Tris HCl, pH 7.9 at 4°C/5mM Mg acetate/10mM mercaptoethanol/50% glycerol). The suspension of nuclei was adjusted to a final concentration of 20 OD_{260nm} (measured in 0.5M NaOH). Digestion of nuclei with DNAase I (PL Chemicals) was performed in the same conditions as used by Wu (7) at 25°C for 10 minutes with varying concentrations of enzyme. The nuclei were digested with S1 nuclease (PL Chemicals) in the same conditions as used by Larsen and Weintraub (28) at 37°C for 30 minutes, supplemented with 5% glycerol. The DNA was extracted from nuclei, whether treated with nuclease or not, by proteinase K treatment, phenol extraction, ethanol precipitation and dialysis. The DNA was kept at 4°C in sterile 10mM Tris HCl, 1mM EDTA, pH 7.5 at 4°C.

Restriction digestion, electrophoresis and blotting

Restriction enzymes were purchased from BRL and DNA was digested in a 5-10-fold excess of restriction enzyme for more than 4 hours in the conditions recommended by the suppliers. The DNA was alcohol-precipitated, redissolved in electrophoresis buffer (40mM Tris acetate, 2mM EDTA) and electrophoretically separated on a 0.8%-1.2% agarose gel, then transferred to nitrocellulose (33) or DBM (31) paper. The desired fragment was purified by gel elution (29) and nick-translated either by the method of Maniatis et al (30) or with a commercial nick-translation kit (Amersham) in the presence of ^{32}P dCTP and ^{32}P dATP (3000 mCi/mmole). The DNA immobilised on the papers was treated with the ^{32}P probes (31) and, after washing, autoradiography was carried out for 2-8 days at -70°C with Kodak XR5 film (Kodak) using an intensifying screen.

RESULTS

HS around the ε-globin gene in K562 cell nuclei

The human K562 cell line synthesizes both embryonic and fetal haemoglobins and, in particular, produces ε- and γ-globin chains (32). Increased amounts of haemoglobin accumulate following induction with factors such as haemin and butyric acid. Allan et al (19,20) have shown that the pattern of transcription of the ε-globin gene in K562 cells is similar to that in normal human embryos.

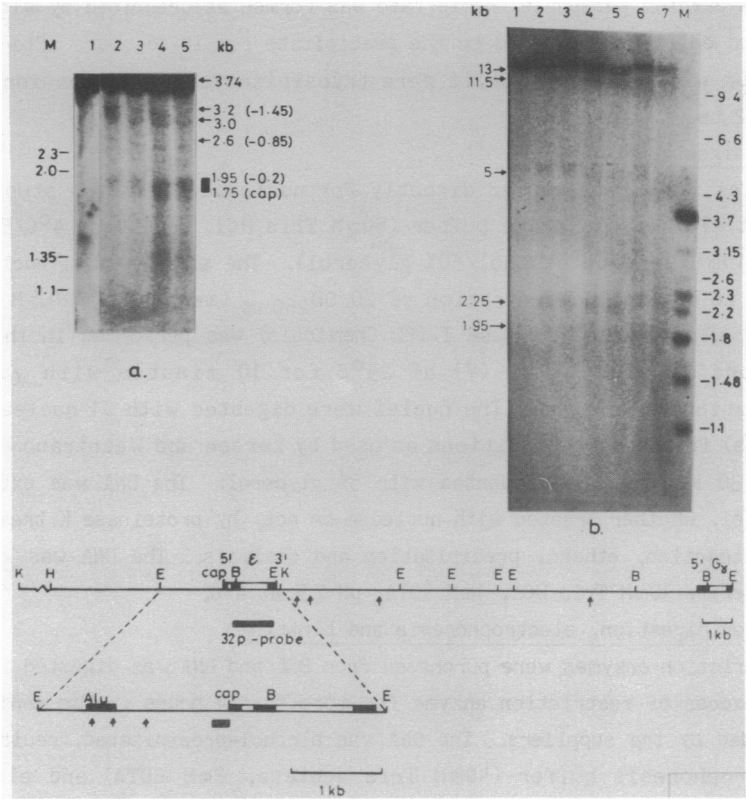

Figure 1

Nuclease Hypersensitive Sites (HS) around the human ε-globin gene in K562 nuclei.

Nuclei were treated with DNAseI or S1 nuclease. DNA was extracted from these and digested to completion with either EcoRI (a) or BamHI (b). The fragments were fractionated by gel electrophoresis, blotted and probed with the ^{32}P labelled BamHI-EcoRI fragment represented in the diagram. HS are indicated by arrows at the side of the autoradiographs and on the line drawing. The sizes of the sub-bands are indicated in kb and the approximate distances of some from the canonical cap site (cap) are shown in brackets. Positions of size markers are indicated in kb in the column M. The locations of restriction sites, ε and $^{G}\gamma$ genes and the more prominent HS are represented in the line drawing (K = KpnI, H = HindIII, E = EcoRI, B = BamHI).

(a) lanes 1-5. K562 nuclei digested with DNAseI at 0, 0.04, 0.06, 0.08, 0.12 µg/ml.

(b) Lanes 1-5. K562 nuclei digested with DNAseI at 0, 0.02, 0.06, 0.12, 0.16 µg/ml.
Lanes 6-7. K562 nuclei digested with 2000 U and 10000 U/ml S1 nuclease.
Lane M. Markers.

In the first series of experiments we studied the distribution of HS around the ε-globin gene in the K562 cell, using the indirect end-labelling technique of Wu (7). Nuclei prepared from induced or uninduced K562 cells were digested with either DNAase I (7) or S1 nuclease (28) and the purified DNA from the nuclease-challenged nuclei was digested to completion with EcoRI or BamHI. The resultant fragments were separated by agarose gel electrophoresis and blotted to either nitrocellulose paper (33) or DBM paper (31). As a probe, we employed a nick-translated 1.275 kb BamHI-EcoRI fragment which contains the second intron, third exon and 154 bp of untranslated 3' flanking sequence and is unique to the ε-globin gene. This provides a convenient probe for both the 3' end of EcoRI fragments and the 5' end of BamHI fragments. These experiments (Figure 1) revealed a region of nuclease hypersensitivity extending from the cap site to 200 bp upstream and three other hypersensitive sites in the 2 kb upstream of the gene. They also revealed several HS within the 4 kb downstream of the gene and a further more distant site 10 kb away (Figure 1). Most of these hypersensitive regions were diffuse rather than discrete sharply localised bands. This is a common observation with vertebrate tissues and implies that the hypersensitive sites represent fairly long regions.

When similar experiments were conducted with nuclei from non-erythroid cells or naked K562 DNA (not shown), no HS were detected in the 5' flanking region of the ε-globin gene. Moreover, we used a probe for the Kappa light chain immunoglobulin gene, which is not transcribed in the K562 cell and, therefore, provides an example of an inactive gene, and found that, when the blots used to demonstrate HS near the ε-globin gene were washed and reprobed with this, no HS could be demonstrated in the region of the gene (results not shown).

These results are in general agreement with many similar reports in the literature which demonstrate HS in and around actively transcribed genes but show either few or no HS in the same genes in cells in which they are not expressed. We were struck by the fact that several of the HS which could be mapped within the 2 kb upstream of the ε-globin gene appeared to correspond to loci which we had previously shown to be initiation sites for minor transcripts. We therefore investigated this relationship in more detail.

First, we confirmed and extended our findings by repeating the experiments, but digesting DNA from nuclease-treated nuclei with KpnI instead of EcoRI. The ε-globin gene is located towards the 3' end of a 15 kb KpnI fragment and the 3' KpnI site is 129 bp downstream from the EcoRI site which

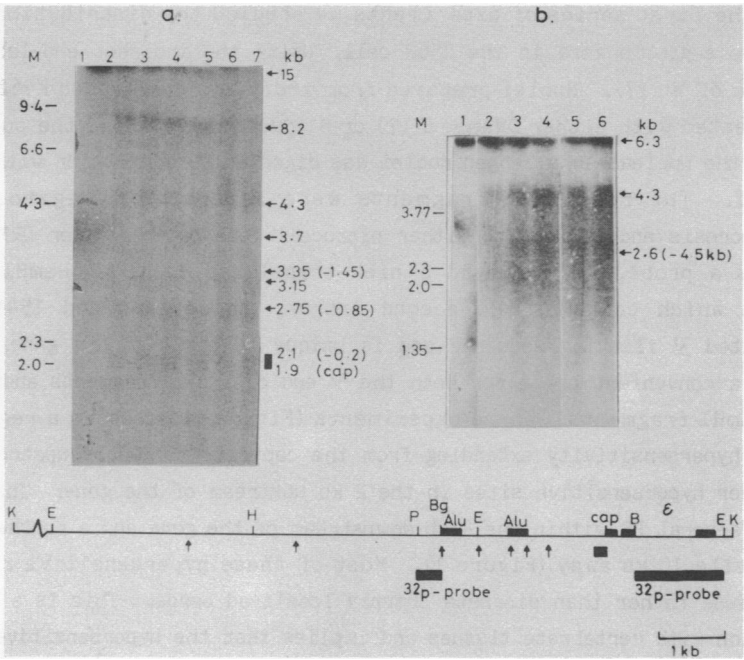

Figure 2

The nuclease hypersensitive sites upstream of the ε-globin gene in K562 nuclei.

K562 nuclei were gently cut with DNAseI (panel a, lanes 2-5, panel b, lanes 4-6) or with S1 nuclease (panel a, lanes 6-7, panel b, lanes 2-3), and DNA from these nuclei was cut to completion with KpnI (panel a) or EcoRI (panel b), separated and blotted on to nitrocellulose paper. DNAseI concentrations were 0, 0.04, 0.08, 0.12 and 0.16 µg/ml for panel a, lanes 1-5 respectively. 0.06, 0.08 and 0.12 µg/ml for panel b, lanes 4-6 respectively. S1 nuclease concentrations were 2000 U and 10000 U/ml for panel a lanes 6 and 7, and panel b lanes 2 and 3 respectively. Lane 1 - no enzyme. Size markers, indicated by dashes at the side of the autoradiographs, were provided by a HindIII digest of λ phage DNA and a HaeIII digest of ϕX174. Nuclease hypersensitive sites are indicated by arrows both on the autoradiographs and the line drawing. The distances in kb of some sites from the canonical cap site are indicated by the figures in brackets. The probe for panel a was the ^{32}P labelled BamHI-EcoRI fragment, and the probe for panel b the ^{32}P labelled PvuII-BglII fragment whose coordinates on the ε-globin gene are shown on the line drawing. (P - PvuII, Bg - BglII, Alu - Alu repeat sequence, other abbreviations as in Fig. 1.)

forms the 3' boundary of the BamHI-EcoRI fragment used as a probe for the previous experiments. Hence, this probe can also be used as an end label for the 15 kb KpnI fragment. As shown in Figure 2a, several HS showed up in this digest. These included the sites demonstrated in the EcoRI fragment which was studied in the previous experiments. It also showed additional sites in

TABLE 1. Nuclease Hypersensitive Sites Upstream of the ε-Globin Gene, Represented by the Distance from the Cap Site in Kilobase Pairs

Initiation sites of transcripts[a] in K562 cells	K562 nuclei DNAseI and S1	B202 nuclei DNAseI and S1	Supercoiled pHR11 DNA S1[b]	Nuclei of Cos 7 transfected pHR11 S1
	-6.3 ± 200			
-4.45	-4.5 ± 200			
	-2.4 ± 200			
	-1.8 ± 200			
-1.43	-1.45 ± 150	-1.4 ± 200	-1.45 ± 150	-1.45 ± 150
	-1.25 ± 150	-1.05 ± 150		
-0.85	-0.85 ± 100	-0.8 ± 100	-0.7 ± 100	-0.85 ± 100
-0.2 to 0[c]	-0.2 to 0[c]	-0.2 to 0[c]	-0.2 to 0[c]	-0.2 to 0[c]

a. Allan et al (20).
b. Allan et al (34).
c. Altogether 8 cap sites occur in this region (see Allan et al (20)).

the adjacent upstream EcoRI fragment, the most distant of which (the 8.2 kb fragment) represents an HS about 6.3 kb upstream of the major cap site.

The location of the further upstream HS was confirmed (Figure 2b) by using a PvuII-BglII probe as an indirect end label for the 6.7 kb EcoRI fragment adjacent to, but upstream of, the EcoRI fragment previously mapped. The location of these HS in relation to the cap site is very consistent from one experiment to another, both in induced and uninduced cells, but the relative density of sub-bands resulting from nuclease cutting varies. For example the site which maps about 4.5 kb upstream from the cap site is barely visible in Fig. 2a (though it can be seen in the original) but it can be seen in Fig. 2b.

The results are summarised in Table 1, alongside the positions of major upstream initiation sites demonstrated by Allan et al (19) for this gene. Two findings are important. The number of HS exceeds the number of demonstrable upstream starts. However, most upstream starts correspond to hypersensitive sites.

HS in cells transfected with recombinant DNA molecules containing the ε-globin gene

When non-erythroid cells such as fibroblasts are transfected with recombinant DNA molecules containing a β-type globin gene, this gene is transcribed although the endogenous gene in the same cell is not transcribed (24). It was, therefore, of interest to investigate the chromatin structure of transferred genes in stably transformed cells. In Figure 3 it is shown that when baby hamster kidney fibroblasts are stably transformed with a

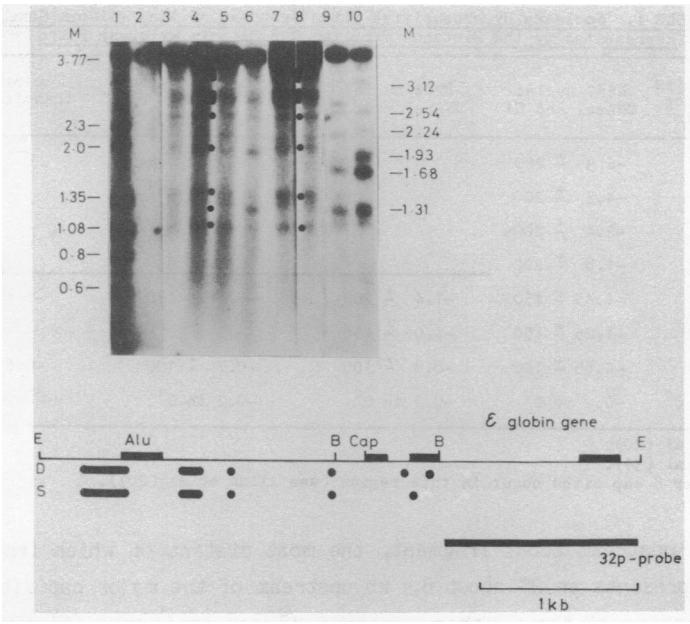

Figure 3

HS in the upstream flanking region of the ε-globin gene in B202 cells (BHK TK$^-$ cells, stably transformed with a recombinant plasmid containing the human ε-globin gene and the HSV thymidine kinase gene).

Nuclei were treated with nucleases and DNA was extracted as described and then digested to completion with EcoRI.

Lane 1 contains markers provided by HindIII digestion of end-labelled λ phage DNA and HaeIII digestion of end-labelled ϕX174. Lanes 2-6 contain DNA from nuclei treated with DNAseI concentrations of 0, 0.1, 0.5, 1 and 2 μg/ml respectively, and lanes 7 and 8 contain DNA from nuclei treated with S1 nuclease at 6000 and 24000 U/ml respectively. The DNA in lane 9 was from nuclei partially digested with AluI (420 U/ml) and in lane 10 from nuclei partially digested with DdeI (480 U/ml) at 37° for 30 minutes. Sub-bands visible in the original, which reflect predominant nuclease cuts in chromatin are indicated by filled circles (DNAseI and S1 nuclease) in both the autoradiograph and the diagram. Since the probe was uniformly labelled sites within the area covered by the probe cannot be mapped accurately; hence, these sub-bands are not indicated in the diagram.

plasmid containing the ε-globin gene, HS are present in the chromatin of this gene and many are in the same locations as in active ε-globin chromatin in K562 cells although additional bands are sometimes seen. Moreover, in transient expression experiments, in which the state of the transfected donor molecules was studied within 48 hours, again NHS were demonstrated not only at the cap site but also at the -0.2, -0.85 and -1.45 loci (Fig. 4), as well as within the gene as shown by smaller sub-bands.

Although there is an overall similarity in the distribution of HS, we

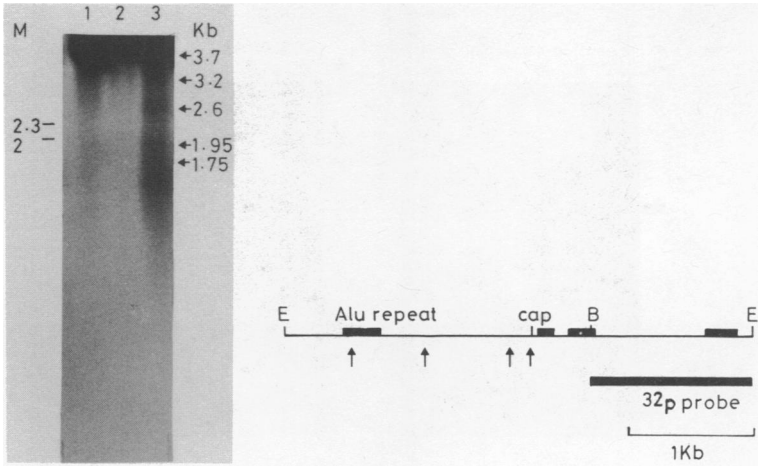

Figure 4

NHS upstream of the ε-globin gene in Cos 7 nuclei, 48 hours after transfection with pHR11.

Nuclei were prepared from Cos 7 cells 48 hours after transfection with pHR11. They were digested with S1 nuclease at a concentration of 100U per 0.02 OD_{260} chromatin at 37° for 30 minutes. The purified DNA from transfected cells (lane 1) and DNA from nuclei from the same cells incubated without (lane 2) or with S1 nuclease (lane 3) were digested with EcoRI. The digests were separated by electrophoresis on 1% agarose gel, blotted to nitrocellulose paper and hybridized to the ^{32}P labelled BamHI-EcoRI fragment (as indicated on the line drawing). The size markers were provided by HindIII digests of λ phage DNA and are shown by arrows at the left side of the autoradiograph. The sub-bands reflecting the S1 nuclease sensitive sites on the ε globin gene are indicated by arrows at the right side of the autoradiograph and on the line drawing.

have observed some differences in individual cases. In K562 nuclei the major HS is diffuse and extends from the region of the major cap site about 200 bp upstream (Figs. 1 and 2). However, in both stably transformed cells (B202, Fig. 3) and short-term transformation experiments (Fig. 4) two separate HS are often demonstrable, one at the canonical cap site and one at -200 bp. The latter is predominant and indeed, in some instances, no HS is visible at the major cap site. These findings may correlate with our observations that the canonical cap site is the major origin of RNA synthesis in K562 cells whereas in the other two instances initiation at the -200 site is much more apparent (34).

These experiments would seem to imply that the chromatin conformation seen when the normal ε-globin gene is actively expressed in erythroid cells is very rapidly assumed when the same gene is introduced to a foreign cell

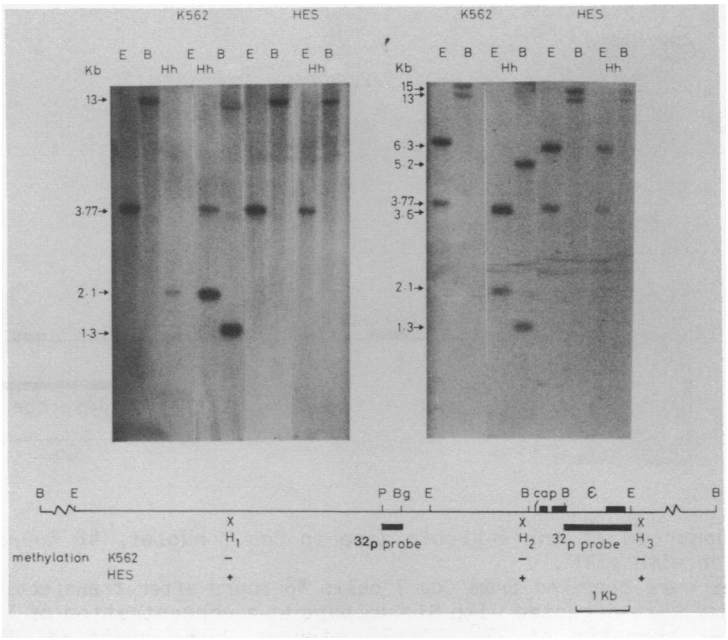

Figure 5

The methylation pattern of the ε-globin gene in K562 cells and HES cells. DNA was cleaved with HhaI (10 units enzyme/1 µg DNA) and subsequently digested with either EcoRI or BamHI. Digested DNA was separated by electrophoresis, blotted on to DBM paper and probed with the ^{32}P labelled BamHI-EcoRI fragment (Fig. 5a). Subsequently the same preparation was also probed with the ^{32}P labelled PvuII-BglII fragment (Fig. 5b). B = BamHI, E = EcoRI, Hh = HhaI, P = PvuII and B = BglII. The location of HhaI recognition sites is indicated on the line drawing at the bottom of the Figure. "+" indicates that the Hha site is methylated, "-" that it is not.

and probably before it becomes integrated into host cell chromosomes. This might be taken to suggest that DNA structure determines the configuration of nucleoprotein of active genes. Indeed, Weintraub (35) has suggested that S1 sensitive sites which are demonstrable in naked supercoiled plasmid DNA have a dominant role in determining the distribution of DNAase I sites in chromatin formed from that DNA. We have made a similar observation. When a supercoiled plasmid, pHR11, which contains the ε-gene linked to pAT153, was investigated for S1 hypersensitive sites using the indirect end-labelling technique, we found that some of the S1 nuclease sensitive sites observed in supercoiled DNA (34) corresponded closely to DNAse I sites in chromatin. The implication of these findings is that DNA structure itself may determine the location of HS. DNAaseI and S1 nuclease cut the chromatin in a similar manner in all instances (Figs. 1, 3 and 4), an observation which is

compatible with single-stranded DNA in HS.

An inverse correlation between the extent of DNA methylation of certain genes and the degree of expression has been established in many viral and nonviral eukaryotic gene systems (36-38). Better agreement has been obtained when methylation of particular CpG sites at the 5' promoter region of the gene is taken into account (39). Hypomethylation and gene activity are also correlated with increased DNAaseI sensitivity and altered chromatin structure of the genes expressed (14,40).

We have compared the methylation state of the ε-globin gene in K562 nuclei with DNA from cell lines such as HES in which the globin gene is not expressed. We digested the total DNA from both cell lines with EcoRI or BamHI supplemented with HhaI which is sensitive to methylation of its recognition site. The separated fragments immobilized to DBM paper were sequentially hybridized either to the ^{32}P labelled BamHI-EcoRI or to the labelled PvuII-BglII fragment. HhaI digestion resulted in cleavage of both the EcoRI and BamHI fragment containing the ε-globin gene in K562 DNA but not in those from HES DNA (Fig. 5). This clearly shows that at least some of the HhaI recognition sites around the ε-globin gene are not methylated in K562 cells but are methylated in HES. This is also true of HpaII recognition sites in the same regions of the ε-globin gene (not shown).

DISCUSSION

HS and promoter function

It has been consistently observed that all actively transcribed genes are characterised by a set of HS which occur in the region of the promoter. Hence, a conventional notion is that the formation of HS at the 5' end of a given gene may be related to promoter function (11). It is striking that we find that the minor initiation sites for the ε-globin gene, as well as the major cap site, map in or near to chromatin regions with exquisite sensitivity to nucleases (Table 1).

Why the correlation occurs is a matter for speculation. The presence of a HS implies a particularly accessible region of DNA and it is possible that if a polymerase II molecule can gain access in many instances it will automatically transcribe. All positions from which transcripts originate correlate fairly closely with the development of HS in the active ε-globin gene but there are some HS for which no such correlation can be seen. In those instances in which no initiation of transcription has been shown to coincide with HS, several explanations are possible. Transcription may occur but at a level too low to detect, the site may already be occupied by another

molecule, the site may not permit access to polymerase II although this is possible for smaller enzymes or the HS may occur at a DNA sequence unsuitable for initiation. However, while the formation of HS may in some instances be a consequence of transcription, there is strong evidence that HS at the major cap site can be formed before the efficient usage of this site for trancription.

The principle has therefore become established that an alteration of chromatin structure is an essential but not sufficient condition for gene transcription; it is based on the following observations: (1) genes can be maintained in an active chromatin configuration after they have been switched off, as judged by an assessment of their overall digestibility and the presence of HS (27); (2) genes can exhibit nuclease hypersensitive sites before they are actually transcribed (4); (3) the transcription frequency is not proportional to the extent of nuclease sensitivity of a given gene.

One particular hypersensitive site may be of special interest. In previous work (20) we obtained strong evidence that the initiation site 4.5 kb upstream of the major cap site represents the 5' limit of transcription. Despite careful study we could find no evidence for RNA synthesis beyond this. However, in the present studies we demonstrated a prominent HS about 6.3 kb upstream. This region presents two striking features. One is a long AT rich region which includes a continuous run of 28Ts. A further 120 bp downstream we have identified a sequence $(AC)_{12}(AT)_6$ followed closely by another more degenerate sequence of alternating purine/pyrimidine nucleotides; this structure potentially can form Z-DNA. We have previously shown that a similar sequence near the mouse β-major globin gene behaves as a negative regulatory sequence (42). This provokes the speculation that structures of this kind may be concerned with regulation of the adjacent gene.

The ε-globin gene is of special interest because not only is it the first β-type globin gene to be expressed in normal ontogeny but it is located at one extreme end of the β-globin cluster. Hence, the upstream boundary of the region could well define the limit of a "domain". The site 6.3 kb upstream may of course be associated with a neighbouring gene, rather than the ε-globin gene and, indeed, the other distant HS, located 10 kb downstream from the gene, may well be related to the $^{G}\gamma$-globin rather than the ε-globin gene.

Two general observations are suggested by these and similar findings from other laboratories. The first is that the functional region or domain associated with an active gene is probably much larger than the sequence

which encodes the major transcript. The major primary transcript of the ε-globin gene arises from 1.5 kb of DNA but our evidence points to an altered conformation of chromatin in the active gene over a region corresponding to at least 11.5 kb of DNA.

Secondly, chromatin apparently adopts a conformation throughout the region bearing hypersensitive sites which is sufficiently highly ordered that landmarks within it can be clearly identified. How does this relate to the "transcriptional unit"? The transcriptional unit is rather loosely defined for eukaryotic genes. We believe it would be useful to define it as that region which contains both sequences coding for RNA precursors and cis-acting structures concerned with control of promoter activity. According to this definition the transcriptional unit for the ε-globin gene extends at least 4.5 kb and probably 6.3 kb upstream of the cap site (20). Our data are not sufficiently complete to postulate a downstream limit.

As this communication was about to be submitted, Tuan and London (43) published a paper describing DNAseI sensitive sites at the 5' side of the ε-globin gene in K562 cells. Many of our findings are in concordance with theirs and discrepancies both in the numbers of sites seen and the positions mapped can quite easily be accounted for by diffuseness of bands and difficulties of standardizing digestion conditions.

ACKNOWLEDGMENTS

We are very grateful for the constructive advice of Dr A Lang and we are indebted to Dr Spandidos for the B202 transformant cell line, to Dr P Anderson for the Kappa light chain probe, to Paul Montague for permitting us to cite his unpublished results and to Dr Oliver Smithies for providing us with sequence data for the upstream region of the ε-globin gene. JDZ was supported by the Chinese Academy of Science, the Henry Lester Trust Ltd and by a British Overseas Research Student Award. The Beatson Institute is supported by the Cancer Research Campaign.

* Permanent Address: Shanghai Institute for Cell Biology, 320 Yo-Yang Road, Shanghai, China

+ To whom correspondence should be addressed

REFERENCES

1. Mathis,D., Oudet,P. and Chambon,P. (1980) Proc. Nucl. Acids Res. Mol. Biol. 24, 1-55.
2. Elgin,S.C.R. (1981) Cell 27, 413-415.
3. Weisbrod,S. (1982) Nature 297, 289-295.

4. Stalder,J., Groudine,M., Dodgson,J.B., Engel,J.D. and Weintraub,H. (1980) Cell 19, 973-980.
5. McGhee,J.D., Wood,W.I., Dolan,M., Engel,J.D. and Felsenfeld,G. (1981) Cell 27, 45-55.
6. Sheffrey,M., Rifkind,P.A. and Marks,P.A. (1982) Proc. Natl. Acad. Sci. USA 79, 1180-1184.
7. Wu,C. (1980) Nature 286, 854-860.
8. Samal,B., Worcel,A., Louis,C. and Schedl,P. (1981) Cell 23, 401-409.
9. Bryan,P.N., Olah,J. and Birnstiel,M.L. (1983) Cell 33, 843-848.
10. Wu,C. and Gilbert,W. (1981) Proc. Natl. Acad. Sci. USA 78, 1577-1580.
11. Sweet,R.W., Chao,M.V. and Axel,R.C. (1982) Cell 31, 347-353.
12. Kuo,M.T., Mandel,J.L. and Chambon,P. (1979) Nucl. Acids Res. 7, 2105-2113.
13. Shermoen,A.W. and Beckendorf,S.K. (1982) Cell 29, 601-607.
14. Burch,J.B. and Weintraub,H. (1983) Cell 33, 65-76.
15. Sargosti,S., Moyne,G. and Yaniv,M. (1980) Cell 20, 65-73.
16. Herbomel,P., Saragosti,S., Blangy,D. and Yaniv,M. (1981) Cell 25, 651-658.
17. McGinnis,W., Shermoen,A.W., Heemsherk,J. and Beckendorf,S.K. (1983) Proc. Natl. Acad. Sci. USA 80, 1065-1067.
18. Emerson,B.M. and Felsenfeld,G. (1984) Proc. Natl. Acad. Sci. USA 81, 95-99.
19. Allan,M., Grindlay,G.J., Stefani,L. and Paul,J. (1982) Nucl. Acids Res. 10, 5133-5147.
20. Allan,M., Lanyon,W.G. and Paul,J. (1983) Cell 35, 187-197.
21. Grindlay,G.J., Lanyon,W.G., Allan,M. and Paul,J. (1984) Nucl. Acids Res. 12, 1801-1820.
22. Carlson,D.P. and Ross,J. (1983) Cell 34, 857-864.
23. Ley,T.J. and Nienhuis,A.W. (1983) Biochem. Biophys. Res. Commun. 112 (3), 1041-1048.
24. Spandidos,D.A. and Paul,J. (1982) EMBO J. 1, 15-20.
25. Pays,E. and Gilmour,R.S. (1981) Biochim. Biophys. Acta 653, 356-367.
26. Wigler,M., Sweet,R., Sim,G.K., Pellicer,A.A., Lacy,E., Maniatis,T., Silverstein,S. and Axel,R. (1979) Cell 16, 777-785.
27. Groudine,M. and Weintraub,H. (1982) Cell 30, 131-139.
28. Larsen,A. and Weintraub,H. (1982) Cell 29, 609-622.
29. Maxam,A.M. and Gilbert,W. (1980) Meth. Enzymol. 65, 499-560.
30. Maniatis,T., Jeffrey,A. and Kleid,D.G. (1975) Proc. Natl. Acad. Sci. USA 72, 1184-1188.
31. Wahl,G., Stern,M. and Stark,G. (1979) Proc. Natl. Acad. Sci. USA 76, 3683-3687.
32. Rutherford,J, Clegg,J.B., Higgs,D.R., Jones,R.N., Thompson,J. and Weatherall,D.J. (1981) Proc. Natl. Acad. Sci. USA 78, 348-352.
33. Southern,E.M. (1975) J. Mol. Biol. 98, 503-517.
34. Allan,M., Zhu,J.D., Montague,P. and Paul,J. (1984) Cell 38, 399-407.
35. Weintraub,H. (1983) Cell 32, 1191-1203.
36. Razin,A. and Riggs,A.D. (1980) Science 210, 604-610.
37. Wigler,M.H. (1981) Cell 24, 285-286.
38. Doerfler,W. (1983) Ann. Rev. Biochem. 52, 92-124.
39. Kruczek,I. and Doerfler,W. (1982) EMBO J. 1, 409-414.
40. Groudine,M., Eisenman,R. and Weintraub,H. (1981) Nature 292, 311-317.
41. Miller,D.M., Turner,P., Nienhuis,A.W., Axelrod,D.E. and Gopalakrishnan,T.V. (1978). Cell 14, 511-521.
42. Gilmour,R.S., Spandidos,D.A., Vass,J.K., Gow,J.W. and Paul,J. (1984). EMBO J. 3, 1263-1272.
43. Tuan,D. and London,I.M. (1984) Proc. Natl. Acad. Sci. USA 81, 2718-2722.