Restriction Endonuclease Accessibility of the Developmentally Regulated Goat γ-, βC-, and βA-Globin Genes in Chromatin

DIFFERENCES IN 5' REGIONS WHICH SHOW UNUSUALLY HIGH SEQUENCE HOMOLOGY*

(Received for publication, April 4, 1984)

Paul A. Liberator‡ and Jerry B. Lingrel§

From the Department of Microbiology and Molecular Genetics, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267-0524

The chromatin structure of the developmentally regulated fetal (γ), preadult (βC), and adult (βA) β-globin genes of the goat has been investigated using a nuclear restriction enzyme accessibility assay. In fetal liver nuclei only the PvuII site immediately proximal to the γ-globin gene is available for digestion, while the homologous recognition sequences 5' to βC and βA are not accessible. Conversely, that site upstream of the βC transcription unit is exclusively digested in bone marrow nuclei prepared from animals expressing the juvenile form of the protein. In nonerythropoietic tissue none of the PvuII recognition sequences flanking the three genes are digested using identical reaction conditions. These results are particularly striking since the nucleotide sequence extending for hundreds of bases in either direction from this restriction site is remarkably homologous among the three genes. In addition, we have mapped an endogenous nuclease-hypersensitive site approximately 1150 nucleotides 5' to the γ-globin gene which is evident only in fetal liver tissue.

Chromosomal structure is necessarily involved in the regulation of eukaryotic gene expression since the majority of nuclear DNA is very highly condensed with proteins and as such is unavailable to the transcriptional machinery. Consequently, that part of the genome which is being actively transcribed exists as chromatin which is structurally distinct from the bulk of nuclear DNA. This was first documented as an increased sensitivity to digestion by DNase I (1, 2). Limited digestion with this endonuclease demonstrated the existence of specific sites in chromatin that are exquisitely sensitive to DNase I (3, 4). Multiple nuclease-hypersensitive sites which correlate with gene expression have now been mapped in various systems where expression is regulated in several distinct ways (reviewed in Ref. 5).

The nature of these hypersensitive sites is not well documented. Electron micrographs of SV40 minichromosomes (6, 7) have revealed a region near the origin of replication that appears to lack a nucleosome particle and which corresponds to a DNase I-sensitive region in the viral genome. The concept of protein-free regions of chromosomal DNA in vivo has been advanced in two studies (8, 9). Both demonstrate a correlation between gene expression, the existence of a DNase I-hypersensitive site, and the ability of appropriate restriction endonucleases to recognize and digest chromatin within the hypersensitive region. Moreover, a 115-bp fragment, located 5' to the chicken βC-globin gene was excised from erythrocyte nuclei by digestion with MspI and appeared to behave electrophoretically as protein-free DNA (8).

In the studies presented here we describe an analysis of the chromatin structure immediately upstream from three of the goat β-like globin genes at two distinct erythropoietic developmental stages. In a pattern similar to that seen in humans, goat globin expression is subject to a series of programmed switches (10, 11). During the first 25 days following conception, embryonic hemoglobin (HbE, αα) is expressed. At approximately this point during gestation, the switch to fetal hemoglobin occurs (HbF, ααγγ), which is followed by a second switch to the preadult form of the protein (HbC, α2βγ) late in fetal life. The final switch to adult hemoglobin (HbA, α2βγ) occurs gradually during the first 6 months postpartum and, unlike other globin developmental switches, is experimentally reversible when the animal is subjected to severe anemia, hypoxia, or erythropoietin treatment (12, 13).

Previous reports from this laboratory have described the isolation from recombinant libraries (14–18) and the characterization (19–22) of 12 genes or pseudogenes in the goat β-globin locus. Although the γ-, βC-, and βA-globin genes are expressed at three separate developmental stages, they are remarkably homologous at the nucleotide sequence level. This sequence homology includes not only the coding regions and intervening sequences (22) but also extends at least 706 bp upstream from the transcription start site (19, 23). Despite the remarkable homology, we demonstrate here that there are distinct differences in the chromatin structure immediately 5' to the three genes when restriction endonuclease accessibility is assayed. These data suggest that the primary sequences which ultimately govern the differential chromatin conformation and perhaps regulate the developmental expres-

* This work was supported by Grants HL 15996 and GM 10999 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by National Institutes of Health Postdoctoral Fellowship 5F32 HL 06416.

§ To whom correspondence and reprint requests should be addressed.

1 The abbreviations used are: bp, base pairs; PMSF, phenylmethylsulfonyl fluoride; DTT, dithiothreitol; SDS, sodium dodecyl sulfate; kb, kilobases; Hb, hemoglobin.
tion of these genes are either very short or are considerably removed from the transcription unit.

EXPERIMENTAL PROCEDURES AND RESULTS

DISCUSSION

Using a restriction enzyme accessibility assay, we have demonstrated that the chromatin fine structure immediately proximal to the goat \( \gamma^- \), \( \beta^- \), and \( \beta^A \)-globin genes changes in a manner which parallels their expression. More specifically, the PvuII recognition sequence located 9 nucleotides upstream from the transcriptional start site in each of the three genes is differentially accessible to digestion (Fig. 5). In nonerythropoietic adult liver nuclei, each of the target recognition sequences is resistant to digestion by the restriction endonuclease. To the contrary, in fetal liver nuclei the PvuII site proximal to the \( \gamma^- \)-globin gene is accessible while the homologous sites adjacent to the \( \beta^- \) and \( \beta^A \) transcription units are not digested. This result indicates that the three genes, or at least their immediate 5' flanking regions, do not share identical chromatin environments in the same tissue. Presumably the \( \beta^- \) and \( \beta^A \)-globin genes, the products of which are not expressed at this developmental stage, are located in a restricted environment, at least by the criterion of PvuII digestion. Bone marrow nuclei prepared from anemic goats with the goat P-like globin gene is preferentially accessible relative to both the \( \gamma^- \) and \( \beta^- \) homologues.

Due to the unique nature of the goat \( \beta^- \)-like globin gene family, this finding is of particular interest. The locus encompasses 12 genes (5'-\( \gamma^-\beta^-\beta^-\beta^-\beta^-\beta^-\beta^-\beta^-\beta^-\beta^-\beta^-\beta^-\beta^-5' \) (18)), spans nearly 120 kb, and probably arose via the triplication of an ancestral 4-gene set (5'-\( \gamma^-\beta^-\beta^-\beta^-5' \)). The \( \gamma^- \), \( \beta^- \), and \( \beta^A \)-globin genes are located in the 3' most position of the three 4-gene blocks. The triplication argument is strengthened by the dramatic similarity in the nucleotide sequence between \( \gamma^- \), \( \beta^- \), and \( \beta^A \). Regions of homology are not restricted to any particular portion of the genes. Instead, there are stretches of identical sequence distributed throughout their length, including both the coding and noncoding regions (22). Furthermore, significant sequence homologies exist a considerable distance away from the transcription unit, especially in the 5' direction (19, 23). This is especially striking in that their expression is subject to strict developmental regulation. Consequently, some significant differences must necessarily exist between the three genes to allow for their appropriate recognition by specific regulatory molecules and differential expression.

Although control of \( \gamma^- \), \( \beta^- \), and \( \beta^A \)-globin gene expression at the transcriptional level has not yet been rigorously demonstrated, the differences recognized in this type of analysis are consistent with that concept. More specifically, it is tempting to draw a parallel between PvuII and RNA polymerase II. For example, in fetal liver nuclei the lack of PvuII digestion proximal to \( \beta^- \) and \( \beta^A \) might be equated to an inability of RNA polymerase II to initiate transcription upon these templates. The observation that the restriction enzyme does cut 5' to the \( \gamma^- \)-globin gene is consistent with its being transcribed in this developmental compartment. This is perhaps an unfair comparison to make for numerous reasons, not the least of which is the marked difference in the molecular weight of the two enzymes. However, the data presented in this paper only address the minimum size of the altered domain and do not reflect the boundaries of the differential chromatin environment. In fact, the length of flanking genomic DNA which is accessible is probably considerably larger. By analogy, McGhee et al. (8) have described a 200-bp region close to the 5' end of the chicken adult \( \beta^- \)-globin gene which is both hypersensitive to DNase I digestion and particularly susceptible to restriction enzymes. Both of these markers for altered chromatin structure were evident only in nuclei isolated from definitive erythrocytes and not in nuclei from primitive erythrocytes which do not express this gene. In a similar type of study, Sweet et al. (9) demonstrated that DNA sequences between positions -4 and -182 from the thymidine kinase gene are both DNase I hypersensitive and accessible to restriction enzymes in a manner which strictly parallels its expression.

It has been proposed that the potential information in DNA secondary structure is used to create hypersensitive sites in chromatin. This proposal is based on the findings that DNase I-hypersensitive regions possess an alternative DNA structure, one that is perhaps transiently single stranded in nature. That is, sites which are hypersensitive to DNase I in chromatin map near areas that are preferentially digested by S1 nuclease (38, 39) and bromoacetaldelyde (40), both of which are single-strand specific reagents. Moreover, these same regions are preferentially sensitive in supercoiled but not in relaxed DNA (38, 41, 42). This occurs in the absence of any chromosomal proteins and implies that certain DNA sequences have a propensity to adopt an altered conformation when placed in the appropriate environment. Furthermore, Weintraub (39) has demonstrated in a series of DNA transfection experiments that the potential to form specific secondary structures in chromatin can be a dominant characteristic.

The results reported here do not contradict this prediction but instead argue for an interaction between distinct chromatin regions. Because of the considerable homology between

---

Fig. 5. Accessibility of the PvuII recognition sequence immediately proximal to the goat \( \beta^- \)-globin genes in nuclei. The figure at the top illustrates a prototype goat \( \beta^- \)-globin gene; open blocks are coding regions, heavy dark lines represent untranslated nucleotide sequence, and CCAAT and ATAAAA depict the location of consensus transcriptional regulatory sequences (19, 22). Symbols used are: \( \downarrow \), PvuII recognition sequence; \( \downarrow \), location of the endogenous nucleasere-hypersensitive site 5' to the \( \gamma^- \)-globin gene in fetal liver nuclei.

- **Tissue Type**: Fetal Liver, Juvenile Bone Marrow, Adult Liver
- **Type of Hemoglobin Expressed**: \( \beta^- \), \( \beta^A \), \( \gamma^- \)
- **PvuII Accessibility**
  - \( \beta^- \): +
  - \( \beta^A \): +
  - \( \gamma^- \): +

---

Portions of this paper (including "Experimental Procedures," "Results," and Figs. 1-4, and Footnote 3) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.
the primary structure of the goat γ-, β′-, and δβ-globin genes and their flanking sequences, it is logical to anticipate that all or none of the three genes has the ability to assume an alternate conformation when packaged into chromatin or supercoiled in a plasmid. Using a restriction enzyme accessibility assay, it is evident (at least for γ and β′) that this in fact is the case. While each sequence has that potential, the three PvuII sites are not similarly accessible within the same tissue. This predics the existence of gene-specific sequences which act in concert with appropriate developmental signals to selectively modify local chromatin structure. As a result of this event, those sequences proximal to the transcription unit which have the propensity to adopt an alternate conformation when packaged into chromatin or tissue. This predicts the existence of gene-specific sequences of the insertion-like elements in their second intervening locus, have moved into a position proximal to the gene.

A final possibility is that sequences which are well removed from the transcription unit ultimately govern the chromatin conformation around these genes and perhaps regulate their developmental expression. The concept of gene regulation by sequences acting at a distance is not without precedence (43-47). One striking example is a human γ-thalassemia which is characterized by a large deletion whose end point maps well upstream from the adult β-globin gene, leaving some 2.5 kb of normal 5′ flanking sequence (48). It appears that either regulatory sequences located greater than 2.5 kb upstream of the transcription unit have been deleted or that actively suppressing sequences, which normally exist outside of the locus, have moved into a position proximal to the gene. Recognizing these examples, it is conceivable that those sequences which act in concert with specific developmental signals to impose the differential chromatin environments flanking the goat γ-, β′-, and δβ-globin genes are also a considerable distance from the respective coding regions.

REFERENCES

9. S. Spence, unpublished observations.

Goat α-Globin Chromatin Structure

15500

Source and Isolation of Nuclei. Goat thymus and bone marrow of 8 to 10-week-old immature females and young adult males were used to prepare bone marrow nuclei, while fetal thymus was obtained from pregnant 100-gram old animals. The type of circulatory system in the goat was determined by observation of the color of the blood photographed under a stereomicroscope on a trinitro-azide-acrylic polymerizing gel (24). Animals were anesthetized with intraperitoneal injections of ketamine hydrochloride and sodium pentobarbital. Blood was collected from the heart by dissection and mice were killed with an intraperitoneal injection of sodium pentobarbital. Thoracic and lumbar vertebrae were removed from anesthetized goats and the heart was excised into a solution containing 0.05 M sodium citrate, 0.1 mg/ml phenylmethylsulfonyl fluoride (PMSF) and aprotinin. Nuclei were isolated as described by Fox et al. (25), briefly, this method involves the digestion of goat thymus nuclei with collagenase enzyme and RNAse A, followed by the addition of trypsin. Trypsin was stopped with EDTA and nuclei were suspended in 0.2 M sucrose, 0.02 M Tris-HCl pH 7.0, 0.02 M PMSF (TPS-150) and 0.02 M EDTA and homogenized in a half-volume of 23 M sucrose containing 0.5 M sucrose using a motor driven Teflon pestle. The homogenate was filtered through cheese cloth and a crude pellet was collected and resuspended in 5 ml of 0.25M sucrose for 5 min. The pellet was then used in the following digestion procedure. The 0.25M sucrose was equivalent to 20 volumes of TMS-150 containing 0.5 M sucrose, passed through cheese cloth and then washed with prewarmed 2.75M sucrose to obtain the final pellet which was used without further digestion. 0.25M sucrose equals 20 volumes of TMS-150 immediately before digestion with restriction enzymes. Nuclear DNA concentration in each solution was determined using the method described by Burton (26). 0.25M sucrose equals 20 volumes of TMS-150 immediately before digestion with restriction enzymes.

Digestion of Nuclei and Isolation of DNA. Fetal (New England Biolabs or American International, Ltd.) digestion was performed at a DNA concentration of 0.2 μg/ml in the following buffer: 10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM MgCl2. Restriction enzyme digestion was performed on nuclei washed with a buffer containing 0.25M sucrose and 0.5M Tris-HCl, pH 8.0, 50 mM NaCl, 5 mM MgCl2, 0.1 mM spermine, 0.1 M spermidine, 0.2M NaCl, 0.2 M sodium acetate, 0.5 M sucrose, and homogenized in 10 volumes of buffer A. Nuclei were homogenized with a Teflon pestle in buffer A containing 0.5 M sucrose using a motor driven Teflon pestle. The homogenate was filtered through cheese cloth and a crude pellet was collected and resuspended in 5 ml of 0.25M sucrose for 5 min. The pellet was then used in the following digestion procedure. The 0.25M sucrose was equivalent to 20 volumes of TMS-150 containing 0.5 M sucrose, passed through cheese cloth and then washed with prewarmed 2.75M sucrose to obtain the final pellet which was used without further digestion. 0.25M sucrose equals 20 volumes of TMS-150 immediately before digestion with restriction enzymes. Nuclear DNA concentration in each solution was determined using the method described by Burton (26). 0.25M sucrose equals 20 volumes of TMS-150 immediately before digestion with restriction enzymes.

Secondary Restriction Digests Blotting and hybridization. Secondary restriction enzyme digestion of the preextracted DNA was used to generate restriction fragments (26-28) of 500 to 5000 base pairs (b.p.) in size. The DNA was sheared to an average size of 500 b.p. by passage through a 22 gauge needle into a 15 ml centrifuge tube. Digestion was performed at a DNA concentration of 0.2 μg/ml in the following buffer: 10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM MgCl2. Restriction enzyme digestion was performed on nuclei washed with a buffer containing 0.25M sucrose and 0.5M Tris-HCl, pH 8.0, 50 mM NaCl, 5 mM MgCl2, 0.1 mM spermine, 0.1 M spermidine, 0.2M NaCl, 0.2 M sodium acetate, 0.5 M sucrose, and homogenized in 10 volumes of buffer A. Nuclei were homogenized with a Teflon pestle in buffer A containing 0.5 M sucrose using a motor driven Teflon pestle. The homogenate was filtered through cheese cloth and a crude pellet was collected and resuspended in 5 ml of 0.25M sucrose for 5 min. The pellet was then used in the following digestion procedure. The 0.25M sucrose was equivalent to 20 volumes of TMS-150 containing 0.5 M sucrose, passed through cheese cloth and then washed with prewarmed 2.75M sucrose to obtain the final pellet which was used without further digestion. 0.25M sucrose equals 20 volumes of TMS-150 immediately before digestion with restriction enzymes. Nuclear DNA concentration in each solution was determined using the method described by Burton (26). 0.25M sucrose equals 20 volumes of TMS-150 immediately before digestion with restriction enzymes.

Results

Nuclei were prepared from the liver of a goat fetus at 70-80 days post-conception and digested with increasing amounts of Eco RI. At this stage of development, the only α-like protein detectable in fetal blood in y-globin. The DNA was purified, digested to completion with Eco RI, fractionated on 0.8% agarose gels, transferred to a membrane support and hybridized to the α4,y-globin probe. Figure 2B shows that as a consequence of this Eco RI digestion, a specific subfragment is generated, apparently in the expense of the 4.8 kb y-globin Eco RI fragment. The sizes of this subfragment corresponds to the predicted 820 bp by Eco RI fragment illustrated in (Fig. 1). The same Eco RI fragment, the intensities of the 4.8 kb α4,y-globin and the 4.8 kb α4,y-globin bands are not diminished, nor is there any indication that the other predicted subfragments are appearing, even with longer exposure times. It might be argued that there is superposition in the fetal liver molecular which is less than developing y-globin. The mouse, which is also developing y-globin, is not apparent due to the apparent Eco RI digestion upstream from the y-globin. Further demonstrate that this is not the case, a parallel 300 bp fetal mouse y-globin gel hybridized to the 64, 86, 100 bp and 100 bp fragments isolated in Fig. 2B. Fig. 3 shows only a single unidentified fragment close to the 64, 86, 100 bp fragment. This is not a single, readily cut site, in concert with a secondary Eco RI digest, the predicted 1200 bp subfragment is distinctly visible. Since it is unlikely that the secondary Eco RI digest actually digested the 4.8 kb α4,y-globin band, it is more likely that the 4.8 kb α4,y-globin band is not present in the fetal mouse liver nuclei, presumably because it is masked in some manner by chromatin proteins. It is conceivable that the predicted subfragments for α4 and y are not appearing due to a restriction site polymorphism in this area of the y-globin repression sequence. To address this possibility, liver DNA from the same fetus was purified, sequentially digested with Eco RI and Hpa II, fractionated on 0.8% agarose gels, transferred to a membrane support, hybridized to the 64, 86, 100 bp and 100 bp and 100 bp fragments isolated in Fig. 2B. The 4.8 kb α4,y-globin DNA sequence is not present due to the presence of a second unidentified 100 bp fragment. This is simply a matter of fact, the prediction of the subfragments is dependent on the size of the unknown fragment. Taken together, these data suggest that sufficient differences exist in the localization of the y-globin and α4,y-globin genes to render their 5′ flanking regions inaccessible to Eco RI digestion. This is especially dramatic in light of the remarkable degree of similarity between various systems for hundreds of nucleotides in either direction from the Eco RI recognition sequence.

Figure 3. Restriction enzyme digests of nucleic acids prepared from goat fetal liver. Nuclei were prepared from the liver of a goat fetus. Nuclei were digested with an increasing amount of Eco RI (26-28), fractionated on 0.8% agarose gels, transferred to a membrane support and probed with radiolabeled α4,y-globin. (A) Nuclei digested with a titration of FIII 0 to 1,000 units of restriction enzyme as described in Experimental Procedures. The large arrow in the left-hand margin indicates the 4.8 kb sub-fragment generated by cleavage of the 64, 86, 100 bp fragment by Eco RI. The second arrow in the left-hand margin indicates the 820 bp subfragment generated hybridization with the α4,y-globin probe. The large arrow in the right-hand margin indicates the 100 bp sub-fragment generated by cleavage of the 4.8 kb fragment. The second arrow in the right-hand margin indicates the 100 bp sub-fragment generated by cleavage of the 4.8 kb fragment. The second arrow in the right-hand margin indicates the 100 bp sub-fragment generated by cleavage of the 4.8 kb fragment.
A tangled feature of the chromatin structure is found in fetal liver nuclei, which appears to be specific for the γ-globin gene. The large arrow in the left margin of Fig. 3C indicates a fragment that is not present in the absence of the γ-globin gene. Under these conditions, the addition of Pvu II to the primary restriction enzyme digest is apparent at a site 5' to the recognition sequence. The γ-globin gene in the mouse, which is capable of directing transcription of the γ-globin gene, also contains this site. In contrast, the non-γ-globin gene in the mouse, which is not capable of directing transcription, contains an additional site that is not present in the mouse. This site is located 5' to the recognition sequence and is therefore not critical for the expression of the γ-globin gene.

Control experiments to help corroborate the differential accessibility to bone marrow nuclei from anemic goats are in fact shown in the right-hand column of Fig. 3. Bone marrow nuclei, spotted with 1 μg of goat clone 9 (a recombinant plasmid containing the entire 5' region of the γ-globin gene) and digested with 500 units of Pvu II, are digested to completion with Eco RI, transferred to non-anemic goats under conditions of high expression of the γ-globin gene, and subjected to a PYU II-Eco RI double digestion. Each of the chromosomes containing the γ-globin gene is digested at that site. Furthermore, when genomic DNA is purified from these nuclei and subjected to a PYU II-Eco RI double digestion, each of the chromosomes containing the γ-globin gene is digested at that site. The data presented in Fig. 3 strongly support the idea that the accessibility of the γ-globin gene to bone marrow nuclei from anemic goats is a consequence of the γ-globin gene and is independent of the presence of the γ-globin gene in the mouse. Therefore, the data presented in Fig. 3 strongly support the idea that the accessibility of the γ-globin gene to bone marrow nuclei from anemic goats is a consequence of the γ-globin gene and is independent of the presence of the γ-globin gene in the mouse.