Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 8953–8957, October 1991
Biochemistry

A DNA-binding factor in adult hematopoietic cells interacts with a pyrimidine-rich domain upstream from the human δ-globin gene

(hemoglobin switching/S1 nuclease-sensitive sites/H-DNA)

DAVID O'NEILL*, KATHERINE BORNSCHLEGEL†, MICHAEL FLAMM‡, MADALYN CASTLE†, AND ARTHUR BANK†‡

Columbia University, College of Physicians and Surgeons, Departments of †Genetics and Development, ‡Medicine, and *Pathology, New York, NY 10032

Communicated by Helen M. Ranney, June 21, 1991 (received for review March 4, 1991)

ABSTRACT **To date, DNA-binding factors with a developmental pattern of expression have not been described in human erythroid cells to explain the switch from fetal (γ-) to adult (δ- and β-) globin gene expression. Here we describe a factor present in nuclear extracts from adult mouse and human hematopoietic cells that binds to an oligopyrimidine repeat approximately 960 base pairs upstream from the human δ-globin gene. The binding site for the factor is within an unusual 250-base-pair domain that is greater than 95% pyrimidines on one strand. This domain is preferentially sensitive to S1 nuclease in supercoiled plasmids, indicating that it can adopt an alternative non-B-DNA conformation. A number of S1-sensitive sites within the domain, including the factor-binding site, have sequence characteristics associated with the formation of a triple helix (H-DNA). The position of the binding site between the fetal and adult β-globin-like genes, its potential for adopting an unusual secondary structure, and the restricted activity of the binding factor to adult hematopoietic tissues suggest possible roles in hematopoietic cell development and hemoglobin switching.**

The human β-globin gene family is a model system to study the regulation of eukaryotic gene expression (1–3). Little is known, however, about the mechanisms by which these genes are switched on and off during development. Understanding this switch could provide insights leading to unique approaches to the treatment of certain inherited human diseases involving hemoglobin, including sickle cell disease and β thalassemia. Analysis of the switch from primitive (embryonic) to definitive (adult) globin synthesis in the chicken indicates that developmental regulation of globin genes is mediated by sequence-specific DNA-binding proteins (4–9). Among the DNA-binding proteins implicated in chicken hemoglobin switching are two factors specific to adult erythroid cells: BGP1, which binds to a poly(dG) sequence in the β-globin promoter (4, 5), and NF-E4, which is believed to mediate the developmental switch by recognizing polypurine sequences within the β-globin promoter and enhancer (8, 9).

Two sequences identical to the chicken NF-E4 binding site are present within the 70-kilobase (kb) human β-globin gene cluster, 5′ and 3′ of the δ-globin gene. Here, we describe a nuclear factor in mouse and human hematopoietic cells that binds to an oligopyrimidine repeat near one of the potential NF-E4 recognition sequences, approximately 960 base pairs (bp) 5′ of the δ-globin gene. Its binding activity is present at high levels in erythroid cells that express adult globin, but only at low levels in cells that express embryonic or fetal globin. Strong binding activity is also found in mouse and human lymphocyte cell lines expressing B- or T-cell surface antigen receptors. The factor is not detected in nonhematopoietic cell lines.

The binding site for this factor is located within an unusual 250-bp domain that is greater than 95% pyrimidines on the coding strand and is preferentially sensitive to S1 nuclease in supercoiled plasmids. S1 nuclease-sensitive polypyrimidine·polypurine [poly(Pyr·Pur)] sequences can adopt a number of non-B-DNA conformations, including a triple helix (H-DNA), which may play an important role in gene regulation (10–20).

MATERIALS AND METHODS

Preparation of Nuclear Extracts. Nuclear extracts from tissue culture cell lines were prepared as previously described (21, 22). MEL and K562 cells were induced to differentiate as described (21, 23). Embryonic blood cells were obtained by bleeding 11-day mouse embryos and their yolk sacs into cold phosphate-buffered saline. Nuclei from mouse 14-day fetal liver and mouse adult liver were isolated according to Gorski *et al.* (24). Nuclear extracts were prepared from these samples as above (22).

DNA Binding Studies. DNase I footprinting and gel mobility shift assays were performed as described (21), except that in gel shift assays electrophoresis was on 4% polyacrylamide gels. The footprinting probe was a 383-bp *Sau*3AI restriction fragment from positions −1223 to −841 relative to the δ-globin cap site (see Fig. 1). This was subcloned in the *Bam*HI site of plasmid pGEM-3Z (Promega) to yield plasmid pGEM-3Z-δ. Double-stranded competitor oligonucleotides were synthesized by Genosys (The Woodlands, TX).

S1 Nuclease Experiments. Supercoiled plasmids were digested with S1 nuclease (Boehringer Mannheim) for 10 min at 37°C in low-salt buffer, as described (12), and analyzed on a 1% agarose gel. Mapping of exact S1 cleavage sites within the upstream δ-globin sequence was carried out by using a modification of the method of Nickol and Felsenfeld (12). Plasmid pGEM-3Z-δ was digested under gentle conditions which resulted in nicked, rather than linear, plasmid DNA (0.05 unit of S1 per μg of DNA). S1-treated DNA was then digested with either *Bam*HI (which removes the *Sau*3AI insert) or *Eco*RI (which cuts once within the insert, see Fig. 1) and 5′-end-labeled with T4 polynucleotide kinase (ref. 25, pp. 10.60–10.61). *Bam*HI end-labeled samples were then secondarily digested with *Eco*RI, and *Eco*RI end-labeled samples were digested with *Bam*HI. This resulted in two S1-treated preparations of a 355-bp *Eco*RI–*Bam*HI upstream δ-globin fragment, one labeled at the 5′ end of the noncoding strand (*Eco*RI site), the other at the 5′ end of the coding strand (upstream *Sau*3AI site). These fragments were purified from 1% agarose gels by elution onto DEAE membranes (ref. 25, pp. 6.24–6.27). Controls not treated with S1 nuclease were similarly prepared. Samples were analyzed on 6% polyacrylamide/8 M urea sequencing gels.

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.

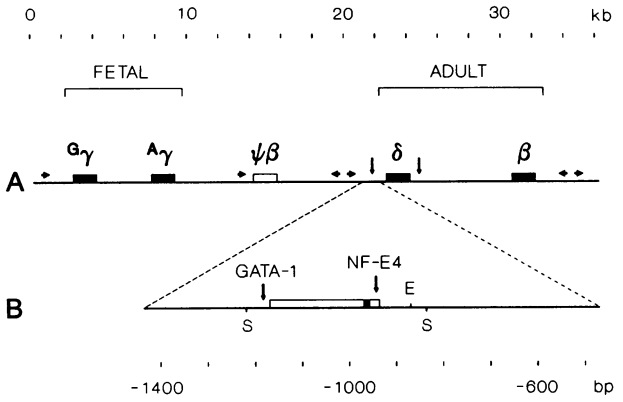

FIG. 1. (*A*) Map of the human fetal and adult β-globin-like genes on chromosome 11. Horizontal arrows mark sites and orientation of *Alu* family sequences. Potential NF-E4 binding sites are indicated by vertical arrows. (*B*) Map of the region of the potential NF-E4 binding site upstream from the δ-globin gene. The open box marks the position of the 250-bp pyrimidine-rich domain. The location of the CCTTCCTTCC binding site is indicated by the filled portion of the box. *Sau*3AI (S) and *Eco*RI (E) sites and potential NF-E4- and GATA-1-binding sites are shown. The scale at the top indicates kilobase distances for map *A*; the scale at the bottom marks positions relative to the δ-globin cap site for map *B*.

RESULTS

Computer Search of the Human β-Globin Gene Cluster. To locate potential NF-E4 recognition sequences within the human β-globin gene cluster, a computer search (26) was done, using RAGAGGRGG (R = any purine) as the query sequence for NF-E4 (8). Interestingly, identical matches to the NF-E4 binding sequence were found only in the vicinity of the δ-globin gene (Fig. 1). The NF-E4 sequences (AAGAGGAGG) are both on the noncoding strand, one 940 bp upstream from the δ-globin transcription initiation site, the other 450 bp downstream from the δ-globin polyadenylylation signal. The upstream NF-E4 sequence lies within an unusual 250-bp tract that is greater than 95% pyrimidines on the coding strand.

DNase I Footprinting. We carried out DNase I footprinting of a *Sau*3AI fragment that contains the entire upstream δ-globin 250-bp pyrimidine-rich sequence and a GATA-1 binding site (Fig. 1) to characterize protein–DNA interactions in this region. Nuclear extracts from mouse erythroleukemia (MEL) cells, which express adult globin, human embryonic/fetal erythroid (K562) cells, mouse T-lymphocytes (EL4), and human nonhematopoietic (HeLa) cells were analyzed. When DNA probe labeled at the 5′ end of the coding (pyrimidine-rich) strand is used, no footprint of the NF-E4 recognition sequence (−935 to −943) is seen (Fig. 2 *Left* and *Center*). By contrast, a clear footprint is seen at the nearby oligopyrimidine repeat CCTTCCTTCC (−958 to −967, referred to hereafter as the *pyr* sequence), but only with extracts from EL4 and uninduced MEL cells (Fig. 2 *Left* and *Center*, lanes 2 and 3, bracket B), and not with extracts from K562, HeLa, or induced MEL cells. A DNase I footprint of the opposite strand shows no protection of either the NF-E4 site or the GGAAGGAAGG sequence complementary to the *pyr* sequence (Fig. 2 *Right*).

A strong footprint, associated with DNase I-hypersensitive sites, is found on both strands at the GATA-1 recognition sequence (−1188 to −1182, Fig. 2 *Left* and *Right*, bracket A) with both uninduced (lanes 3 and 5) and induced (lanes 4 and 6) erythroid nuclear extracts. A weak footprint covering the GATA-1 sequence is also seen with HeLa nuclear extracts.

Gel Mobility Shift Assays. The footprinting data suggest that the activity of the *pyr* sequence-binding factor is restricted in both a tissue-specific and developmental stage-specific manner, since it is detected in nuclear extracts from

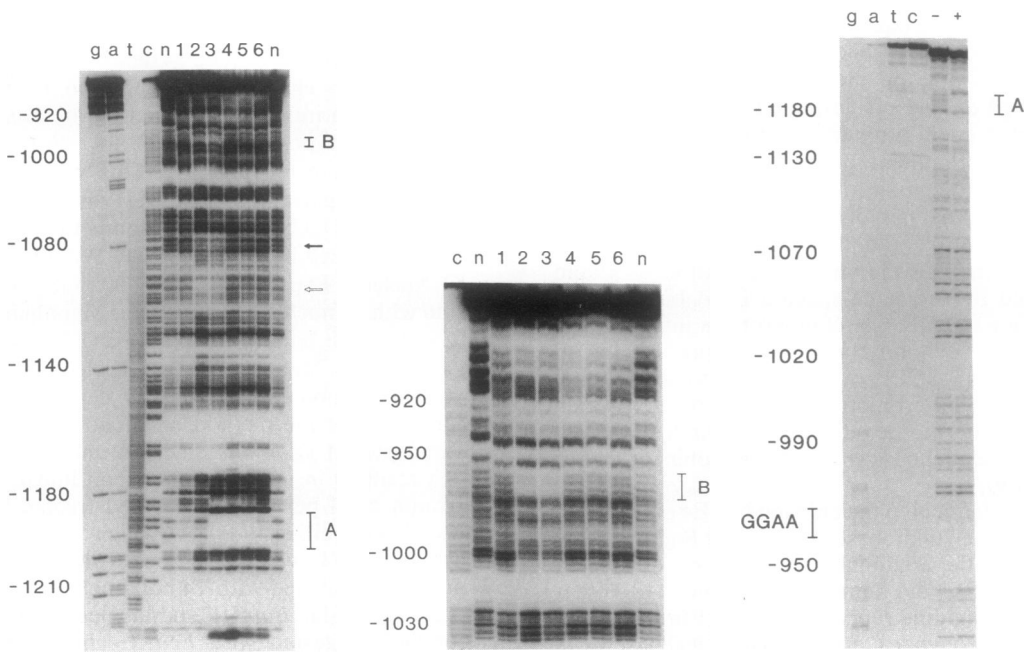

FIG. 2. DNase I footprint analysis of the −1223 to −841 upstream δ-globin region. (*Left*) *Sau*3AI probe labeled at the 5′ end of the coding (pyrimidine-rich) strand. Lanes g–c, G, A+G, T+C, and C Maxam–Gilbert sequencing reactions, respectively. Lanes n, control reactions with no nuclear extract. Lane 1, HeLa; lane 2, EL4; lane 3, uninduced MEL; lane 4, induced MEL; lane 5, uninduced K562; lane 6, induced K562. Brackets A and B indicate sites protected from DNase I: A marks the GATA-1 footprint; B marks the footprint at the CCTTCCTTCC (*pyr*) sequence. Arrows mark areas of weak protection with MEL and EL4 nuclear extracts centered on positions −1069 (filled arrow) and −1094 (open arrow). The scale on the left marks positions relative to the δ-globin cap site. (*Center*) Footprint identical to that in *Left*, but run further showing better detail at sites far from the labeled 5′ end. (*Right*) DNase I footprint of the *Sau*3AI probe labeled at the 5′ end of the noncoding (purine-rich) strand. Lane −, no nuclear extract; lane +, MEL nuclear extract. Bracket A marks the GATA-1 binding site protected with MEL nuclear extract. Bracket GGAA marks the GGAAGGAAGG sequence complementary to the *pyr* sequence footprint.

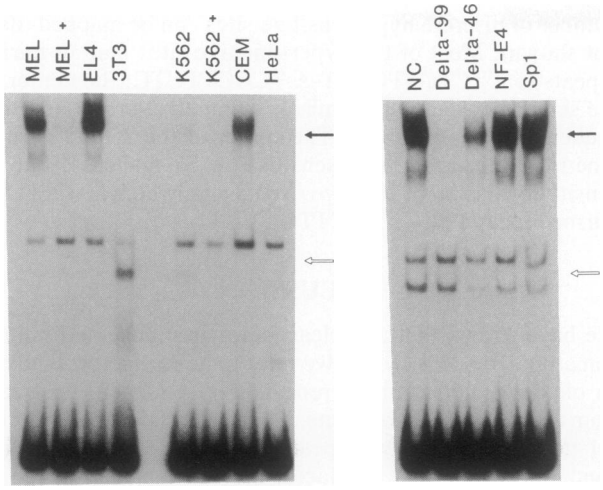

FIG. 3. Gel mobility shift assays using a *Fok* I–*Nla* III restriction fragment probe (δ-99) containing the CCTTCCTTCC binding site. (*Left*) Labeled probe was incubated with nuclear extracts from MEL, induced MEL (MEL +), EL4, 3T3, K562, induced K562 (K562 +), CEM, and HeLa cell lines. The filled arrow marks a slow-migrating protein–DNA complex with a pattern of tissue distribution similar to the *pyr* sequence footprint in Fig. 2. (*Right*) Competition gel mobility shift assays, MEL nuclear extract. Lane NC is a control reaction with no competitor DNA. Approximately 100-fold molar excess of unlabeled δ-99 probe or the following double-stranded competitor oligonucleotides were added to the DNA-binding reaction: δ-46 (5′-TTCCATCCTCTCTCTTCCCCTCTTCCTTCCTTCCTTTCTC-CATTTC-3′), containing the upstream δ-globin *pyr* sequence; NF-E4 (5′-AATAGCTTCCTCCTCTTACTTGCCCCAG-3′), containing the potential NF-E4 recognition sequence at the 3′ end of the human δ-globin gene; and Sp1 (5′-TCGATGGGCGGAGTTAGGGGCGG-GAC-3′), from the simian virus 40 early promoter (27). The open arrow designates bands of higher mobility with which unlabeled δ-99 probe does not compete.

T lymphocytes (EL4) and adult erythroid (MEL) cells, but not in nonhematopoietic (HeLa) or embryonic/fetal erythroid (K562) nuclear extracts. To confirm this, nuclear extracts from these and other cell lines were analyzed by using gel mobility shift assays. Initial experiments using a 46-bp oligonucleotide probe (δ-46) spanning positions −946 to −991 (which includes the *pyr* sequence but not the potential NF-E4 site; see Fig. 1) detected no binding activity with the pattern of the *pyr* sequence footprint (data not shown). This suggested that additional sequence was needed for the factor to recognize the probe, or that the longer probe used for footprinting could adopt a secondary structure needed for the factor to bind which the shorter δ-46 probe could not. To test these possibilities a 99-bp *Fok* I–*Nla* III restriction fragment (δ-99) spanning nucleotides −1001 to −903 was used, which includes the *pyr* sequence and the potential NF-E4 site. With this probe a dense, slowly migrating band is observed with nuclear extracts from uninduced MEL and EL4 cells (Fig. 3 *Left*, filled arrow); only a faint band with the same mobility is seen when uninduced K562 and induced MEL nuclear extracts are used. The band is not detected when HeLa or induced K562 nuclear extracts are used. The DNA-binding activity represented by this band correlates well with the *pyr* sequence-binding activity in DNase I footprints (Fig. 2 *Center*). The decrease observed in *pyr* sequence-binding activity in MEL cells that have been induced to differentiate (Figs. 2 *Center* and 3 *Left*) may indicate that the binding factor disappears upon erythroid differentiation, but it could also be a more direct result of treatment with dimethyl sulfoxide.

Nuclear extracts from human T-lymphocyte (CEM) and mouse fibroblast (3T3) cell lines were also tested (Fig. 3 *Left*). CEM extracts show binding activity comparable to that of EL4 and uninduced MEL extracts, while 3T3 extracts demonstrate no binding activity. Extracts from human B lymphocytes (Daudi) show binding activity similar to that of the T-cell lines (data not shown). These data indicate that strong *pyr* sequence-binding activity is restricted to hematopoietic cells in both mice and humans.

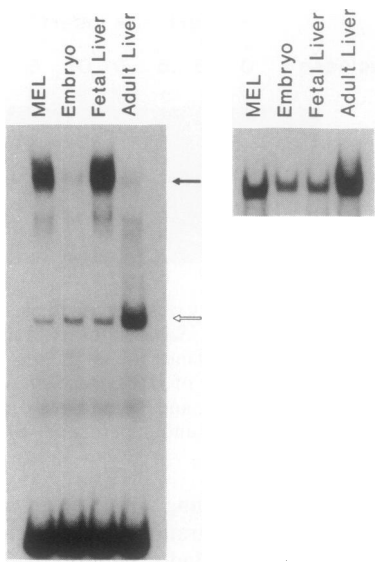

FIG. 4. (*Left*) Gel mobility shift assays using nuclear extracts from mouse hematopoietic tissues and δ-99 probe. *pyr* sequence-binding activity (filled arrow) is present at much higher levels in mouse fetal liver than in embryonic/yolk sac blood cells (see text). The higher-mobility bands in all lanes (open arrow) represent nonspecific binding activity. (*Right*) Control gel mobility shift assays using a human $^{A}\gamma$-globin promoter region probe (21) which binds the ubiquitous octamer-binding protein OTF-1 (unbound probe not shown). The 11- (embryo) and 14-day (fetal liver) mouse extracts demonstrate comparable OTF-1 binding activities.

Competition gel shift experiments were done to determine the sequence specificity of the *pyr* sequence-binding factor. As seen in Fig. 3 *Right*, unlabeled δ-99 inhibits formation of the slow-migrating complex (filled arrow) but not the higher-mobility bands (open arrow), indicating that only the slow-migrating complex represents sequence-specific binding. The shorter δ-46 oligonucleotide also inhibits formation of the slow-migrating complex, but much less effectively than δ-99. Oligonucleotides containing NF-E4 or Sp1 recognition sequences do not compete. These results suggest that the *pyr* sequence-binding factor specifically recognizes the immediate region of the *pyr* sequence repeat, but the presence of the extra flanking sequence in the δ-99 probe significantly enhances binding. The NF-E4 sequence alone is not recognized by the factor.

***pyr* Sequence-Binding Activity in Mouse Hematopoietic Tissues.** To determine if the high level of *pyr* sequence-binding activity in MEL cells compared with K562 cells represents a developmental stage difference *in vivo*, nuclear extracts from mouse hematopoietic tissues at different stages of development were analyzed by gel shift assays (Fig. 4). Strong binding activity is seen in extract from 14-day mouse fetal liver, which expresses only adult globin (Fig. 4 *Left*). By contrast, only faint binding activity is seen in extract from 11-day embryo and yolk sac blood cells, which express only embryonic globin. Binding activity is very weak in adult liver, and may be due to contaminating lymphocytes. These data indicate that strong *pyr* sequence-binding activity *in vivo* is present in adult hematopoietic tissue but not in embryonic hematopoietic or adult nonhematopoietic cells.

S1 Nuclease Sensitivity of the 5′ δ-Globin Pyrimidine-Rich Domain. DNA sequences with predominantly pyrimidines on one strand and purines on the other [poly(Pyr·Pur) se-

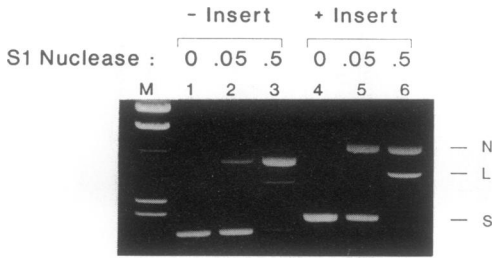

FIG. 5. S1 nuclease sensitivity of supercoiled plasmid DNA with and without upstream δ-globin *Sau*3AI insert. Lane M, marker *Hin*dIII-digested λ phage DNA; lanes 1–3, pGEM-3Z treated with 0, 0.05, and 0.5 unit of S1 per μg of DNA, respectively; lanes 4–6, pGEM-3Z-δ treated with 0, 0.05, and 0.5 unit of S1 per μg of DNA. Nicked circular (N), linear (L), and supercoiled (S) plasmid DNA conformations are indicated.

quences] are capable of forming a number of unusual structures sensitive to the single strand-specific nuclease S1, both in chromatin and in supercoiled plasmids (10–20). To determine if the upstream δ-globin pyrimidine-rich domain is preferentially sensitive to S1 nuclease, we compared the relative S1 sensitivity of plasmids pGEM-3Z-δ, and pGEM-3Z. Fig. 5 shows that pGEM-3Z-δ is noticeably more sensitive to S1 digestion, demonstrated by a more rapid accumulation of linear DNA (L), and a more rapid disappearance of the supercoiled form (S) with increasing amounts of S1.

To map the specific sites of S1 cleavage within the 5′ δ-globin insert, S1-treated pGEM-3Z-δ was radioactively labeled and analyzed on sequencing gels. Comparing DNA treated with S1 nuclease (Fig. 6 *Upper*, lanes 1 and 5) with untreated DNA (Fig. 6 *Upper*, lanes 2 and 6), S1 nuclease-hypersensitive regions are seen on both strands, beginning approximately 100 bp from either 5′ end of the δ-globin insert (lanes 1 and 5, bracket HS). When the gel is run further, a number of discrete hypersensitive sites can be mapped (data not shown). Most of the hypersensitive sites map to mirror repeats (e.g., CTCTTCCCT-A-TCCCTTCTC; the center of the sequence is an axis of mirror symmetry) or poly(dT·dA) sequences (Fig. 6 *Lower*). The center of the 250-bp pyrimidine-rich domain is most sensitive to S1 nuclease, but S1 sensitivity also maps to the *pyr* sequence, which is within the mirror repeat TTCCTTC-CTTCCTT.

DISCUSSION

We have shown that a nuclear factor in mouse and human hematopoietic cells, which we refer to as *pyr* factor, binds to an oligopyrimidine mirror repeat (*pyr* sequence) upstream from the human δ-globin gene. The factor/DNA complex on gel mobility shift assays appears as a dense slowly migrating band, suggesting that this factor may be unusually large or composed of multiple subunits. DNA-binding studies using extracts from erythroid cell lines and mouse tissue indicate that *pyr* factor binding activity is restricted to hematopoietic tissues at late stages in development.

The *pyr* sequence is within a highly unusual 250-bp pyrimidine-rich domain located between the fetal (γ-) and adult (δ- and β-) globin genes. Poly(Pyr·Pur) sequences are capable of forming unusual structures in chromatin and supercoiled plasmids that are preferentially sensitive to the single strand-specific nuclease, S1 (10–20). These structures are thought to play an important role in eukaryotic gene regulation (16). We demonstrate that the 250-bp upstream δ-globin domain is sensitive to S1 nuclease in supercoiled plasmids, and we map a number of discrete hypersensitive sites within this region to mirror repeats (including the *pyr* sequence) and poly(dT·dA) sequences. Interestingly, poly(Pyr·Pur) mirror repeats are known to form an intramolecular triple helix (H-DNA) under the appropriate conditions (15–20).

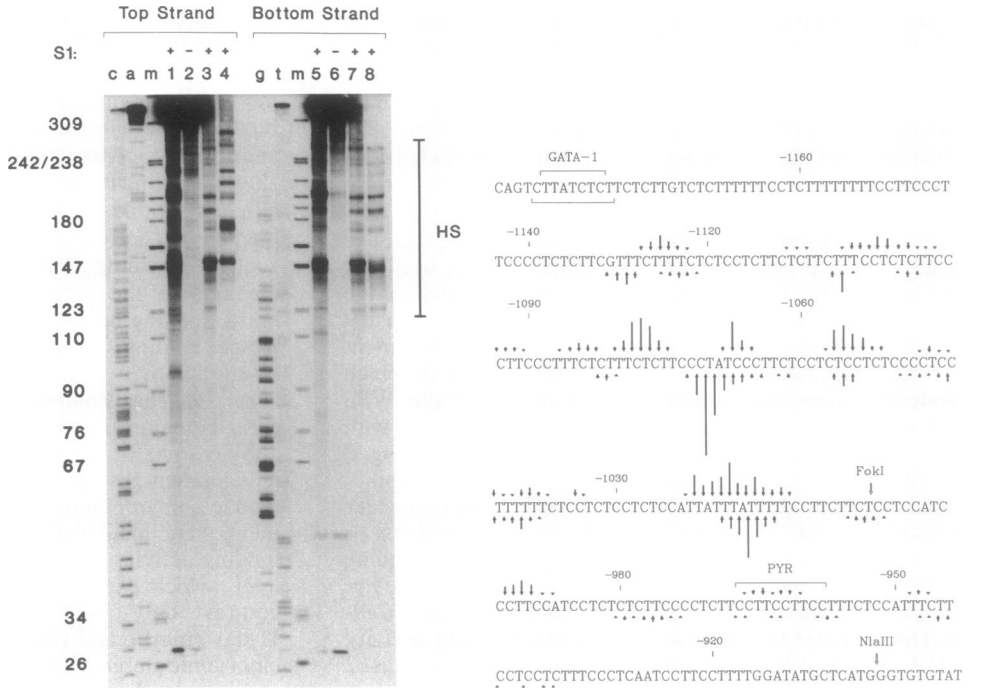

FIG. 6. Mapping of S1 nuclease-sensitive sites within the upstream δ-globin pyrimidine-rich domain. (*Upper*) Reactions on left show DNA 5′-end-labeled on the coding (top) strand; reactions on right shown DNA 5′-end-labeled on the noncoding (bottom) strand. Lanes c, a, g, and t, Maxam–Gilbert sequencing reactions; lanes m, pBR322/*Msp* I marker DNA (size of marker bands are shown on left in bp). Lanes 1 and 5, S1 nuclease-treated DNA; lanes 2 and 6, controls not treated with S1 nuclease. S1 nuclease-hypersensitive regions in lanes 1 and 5 are designated by bracket HS. Some bands of identical mobility in lanes 1 and 5 (e.g., position 149) represent DNA labeled at S1 cleavage sites and are indicated by the control reaction in lanes 3 and 7. In this reaction, DNA samples were end-labeled after treatment with S1 but prior to the *Bam*HI/*Eco*RI restriction digestions. Lanes 4 and 8 are similarly labeled controls digested with *Bam*HI alone (lane 4) or *Eco*RI alone (lane 8), which determine the DNA strand labeled at a particular S1-sensitive site: identical bands in lanes 3 and 4 represent S1 cleavage sites labeled on the noncoding strand, while identical bands in lanes 7 and 8 represent S1 cleavage sites labeled on the coding strand. (*Lower*) Summary of DNase I footprinting and S1 nuclease mapping experiments. The coding (pyrimidine-rich) strand is shown. Nucleotide positions are numbered relative to the δ-globin cap site. Brackets indicate GATA-1 and *pyr* sequence footprints seen in Fig. 2. Arrows or brackets above the DNA sequence represent S1 cleavage or DNase I protection, respectively, on the coding strand; symbols below the sequence refer to the noncoding strand. The degree of S1 sensitivity, as determined by densitometric scanning, is indicated by arrow length.

S1 nuclease-sensitive poly(Pyr·Pur) sequences within promoter elements of a number of genes have been shown to bind nuclear proteins, and are thought to function in transcriptional regulation (28–32). The binding sites for the developmental stage-specific erythroid DNA-binding proteins of the chicken, NF-E4 and BGP1, are also S1 nuclease-sensitive poly(Pyr·Pur) sequences (4, 5, 8, 12, 13). Such DNA sequences exist in equilibrium between different conformations, depending on environmental conditions (17, 20, 28, 33, 34). An important function of factor binding could be to stabilize a particular B- or non-B-DNA structure (17). In the case of hemoglobin switching, the interaction of *pyr* factor with its upstream δ-globin binding site could possibly induce conformational changes within the β-globin gene cluster late in erythroid development. DNase I footprints at this site show protection of only the polypyrimidine strand (Fig. 2), which would be consistent with factor bound to either the three-stranded segment or the nonpaired pyrimidine loop of the H-DNA conformation (18). The dependence of factor binding in gel shift assays on the length of the DNA probe also suggests that *pyr* factor recognizes DNA with an unusual secondary structure; others have shown similar length dependence in the cleavage of linear poly(Pyr·Pur) sequences by S1 nuclease (35, 36).

Interestingly, a sequence highly homologous to the *pyr* sequence is present 2.8 kb upstream from the mouse β-minor globin gene, within a 181-bp homopyrimidine tract (37). A 41-bp oligonucleotide containing this sequence competes with *pyr* sequence-binding activity on gel shift assays (data not shown). This evolutionary conservation suggests a function in globin gene regulation. In addition, studies of natural mutations in humans have suggested that the region between the fetal and adult β-globin-like genes containing the *pyr* sequence may have regulatory importance, particularly in the repression of γ-globin synthesis in adults (1, 2, 38, 39). Experiments using transgenic mice, however, have so far yielded conflicting evidence on the importance of this region in the down-regulation of fetal globin genes (40–42).

It has been proposed that families of short repetitive DNA sequences might provide the structural basis for the coordinate induction of genes in the process of development (43). Interestingly, the CCTTCCTTCC sequence occurs at two additional sites within the 70-kb human β-globin gene cluster, in the first introns of the $^{G}\gamma$- and $^{A}\gamma$-globin genes. If *pyr* factor recognizes these sequences, binding at these sites could block transcription of γ-globin mRNA. Premature termination of transcription at a specific sequence is a well-characterized mechanism for the down-regulation of some genes (44). In a number of instances, this transcription arrest is thought to be mediated through the formation of unusual secondary structures, involving either the template DNA or the transcribed RNA (45). The transcription elongation block in these cases may be modulated by DNA- or RNA-binding proteins (45).

The restriction of *pyr* factor activity to hematopoietic tissues at late stages in development, the location of its DNA-binding site between the fetal and adult globin genes and the ability of the binding site to adopt a non-B-DNA conformation suggest that the factor may function in hematopoietic cell development and, in erythroid cells, hemoglobin switching. The characterization of *pyr* factor and its interactions with the upstream δ-globin and potential γ-globin binding sequences should help clarify its possible function in globin gene regulation.

We are grateful to Larysa Pevny and Dr. Frank Costantini for technical assistance, and to Drs. Costantini, Kathryn Calame, and Argiris Efstratiadis for their comments on the manuscript. This work was supported by grants from the National Institutes of Health (AM07373, DK25274, HL07838, and HL28381) and by the March of Dimes–Birth Defects Foundation.

1. Bank, A., Mears, J. G. & Ramirez, F. (1980) *Science* **207**, 486–493.
2. Stamatoyannopoulos, G. & Nienhuis, A. W. (1987) in *The Molecular Basis of Blood Diseases*, eds. Stamatoyannopoulos, G., Nienhuis, A. W., Leder, P. & Majerus, P. W. (Saunders, Philadelphia), pp. 66–105.
3. Orkin, S. H. (1990) *Cell* **63**, 665–672.
4. Lewis, C. D., Clark, S. P., Felsenfeld, G. & Gould, H. (1988) *Genes Dev.* **2**, 863–873.
5. Clark, S. P., Lewis, C. D. & Felsenfeld, G. (1990) *Nucleic Acids Res.* **18**, 5119–5126.
6. Emerson, B. M., Nickol, J. M. & Fong, T. C. (1989) *Cell* **57**, 1189–1200.
7. Jackson, P. D., Evans, T., Nickol, J. M. & Felsenfeld, G. (1989) *Genes Dev.* **3**, 1860–1873.
8. Gallarda, J. L., Foley, K. P., Yang, Z. & Engel, J. D. (1989) *Genes Dev.* **3**, 1845–1859.
9. Choi, O.-R. & Engel, J. D. (1988) *Cell* **55**, 17–26.
10. Larsen, A. & Weintraub, H. (1982) *Cell* **29**, 609–622.
11. Hentschel, C. C. (1982) *Nature (London)* **295**, 714–716.
12. Nickol, J. M. & Felsenfeld, G. (1983) *Cell* **35**, 467–477.
13. Schon, E., Evans, T., Welsh, J. & Efstratiadis, A. (1983) *Cell* **35**, 837–848.
14. Mace, H. A. F., Pelham, H. R. B. & Travers, A. A. (1983) *Nature (London)* **304**, 555–557.
15. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A. & Frank-Kamenetskii, M. D. (1987) *Nature (London)* **330**, 495–497.
16. Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M. & Wohlrab, F. (1988) *FASEB J.* **2**, 2939–2949.
17. Shimizu, M., Hanvey, J. C. & Wells, R. D. (1990) *Biochemistry* **29**, 4704–4713.
18. Hanvey, J. C., Shimizu, M. & Wells, R. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6292–6296.
19. Voloshin, O. N., Mirkin, S. M., Lyamichev, V. I., Belotserkovskii, B. P. & Frank-Kamenetskii, M. D. (1988) *Nature (London)* **333**, 475–476.
20. Htun, H. & Dahlberg, J. E. (1988) *Science* **241**, 1791–1796.
21. O'Neill, D., Kaysen, J., Donovan-Peluso, M., Castle, M. & Bank, A. (1990) *Nucleic Acids Res.* **18**, 1977–1982.
22. Schreiber, E., Matthias, P., Muller, M. M. & Schaffner, W. (1990) *Nucleic Acids Res.* **17**, 6419.
23. LaFlamme, S., Acuto, S., Markowitz, D., Vick, L., Landschultz, W. & Bank, A. (1987) *J. Biol. Chem.* **262**, 4819–4826.
24. Gorski, K., Carneiro, M. & Schibler, U. (1986) *Cell* **47**, 767–776.
25. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
26. Devereaux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
27. Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tjian, R. (1987) *Cell* **51**, 1079–1090.
28. Boles, T. C. & Hogan, M. E. (1987) *Biochemistry* **26**, 367–376.
29. Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. & Hogan, M. E. (1988) *Science* **241**, 456–459.
30. Postel, E. H., Mango, S. E. & Flint, S. J. (1989) *Mol. Cell. Biol.* **9**, 5123–5133.
31. Johnson, A. C., Jinno, Y. & Merlino, G. T. (1988) *Mol. Cell. Biol.* **8**, 4174–4184.
32. Hoffman, E. K., Trusko, S. P., Murphy, M. & George, D. L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2705–2709.
33. Kohwi, Y. & Kohwi-Shigematsu, T. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3781–3785.
34. Fowler, R. F. & Skinner, D. M. (1986) *J. Biol. Chem.* **261**, 8994–9001.
35. Htun, H., Lund, E. & Dahlberg, J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7288–7292.
36. Evans, T. & Efstratiadis, A. (1986) *J. Biol. Chem.* **261**, 14771–14780.
37. Shehee, W. R., Loeb, D. D., Adey, N. B., Burton, F. H., Casavant, N. C., Cole, P., Davies, C. J., McGraw, R. A., Schichman, S. A., Severynse, D. M., Voliva, C. F., Weyter, F. W., Wisely, G. B., Edgell, M. H. & Hutchison, C. A. (1989) *J. Mol. Biol.* **205**, 41–62.
38. Mears, J. G., Ramirez, F., Leibowitz, D. & Bank, A. (1978) *Cell* **15**, 15–23.
39. Maeda, N., Bliska, J. B. & Smithies, O. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5012–5016.
40. Raich, N., Enver, T., Papayannopoulou, T., Costantini, F. & Stamatoyannopoulos, G. (1990) *Blood* **76**, Suppl. 1, 74 (abstr.).
41. Behringer, R. R., Ryan, T. M., Palmiter, R. D., Brinster, R. L. & Townes, T. M. (1990) *Genes Dev.* **4**, 380–389.
42. Dillon, N. & Grosveld, F. (1991) *Nature (London)* **350**, 252–254.
43. Davidson, E. H., Jacobs, H. T. & Britten, R. J. (1983) *Nature (London)* **301**, 468–470.
44. Spencer, C. A. & Groudine, M. (1990) *Oncogene* **5**, 777–785.
45. Mechti, N., Piechaczyk, M., Blanchard, J.-M., Jeanteur, P. & Lebleu, B. (1991) *Mol. Cell. Biol.* **11**, 2832–2841.