MOLECULAR AND CELLULAR BIOLOGY, Jan. 1989, p. 241–251
0270-7306/89/010241-11$02.00/0

Vol. 9, No. 1

Transcriptional Promiscuity of the Human α-Globin Gene

E. WHITELAW, P. HOGBEN, O. HANSCOMBE, AND N. J. PROUDFOOT*

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, United Kingdom

Received 27 June 1988/Accepted 26 September 1988

The human α-globin gene displays the unusual property of transcriptional promiscuity: that is, it functions in the absence of an enhancer when transfected into nonerythroid cell lines. It is also unusual in that its promoter region lies in a hypomethylated *Hpa*II tiny fragment (HTF) island containing multiple copies of the consensus sequence for the SP1-binding site. We have investigated whether there is a relationship between these two observations. First, we investigated the mouse α-globin gene since it does not lie in an HTF island. We have demonstrated that it was not transcriptionally promiscuous. Second, we studied the transcriptional activity of the human α-globin gene in the absence of the GC-rich region containing putative SP1-binding sites and found a small (two- to threefold) but consistent positive effect of this region on transcriptional activity in both nonerythroid and erythroid cell lines. However, this effect did not account for the promiscuous nature of the human α-globin gene. We found that in a nonreplicating system, the human α-globin gene, like that of the mouse, required a simian virus 40 enhancer in order to be transcriptionally active in nonerythroid and erythroid cell lines. Since we only observed enhancer independence of the human α-globin gene in a high-copy-number replicating system, we suggest that competition for *trans*-acting factors could explain these results. Finally, our experiments with the erythroid cell line Putko suggest that there are no tissue-specific enhancers within 1 kilobase 5′ of the human α-globin cap site or within the gene itself.

The human α-globin gene is unusual since it is transcriptionally active in nonerythroid cell lines in the absence of a simian virus 40 (SV40) enhancer (21, 31, 38). Most tissue-specific genes are not capable of transcriptional activity in heterologous cell lines unless they are placed relatively close to an SV40 enhancer, e.g., human β-globin and δ-globin (21) and rabbit β-globin (3). The DNA at the 5′ end of the human α-globin gene is also unusual in that it is highly GC rich and forms an *Hpa*II tiny fragment (HTF) island. HTF islands are DNA sequences in which CpG is abundant and nonmethylated, resulting in the presence of frequent *Hpa*II restriction sites (8). The human α-globin CpG cluster is nonmethylated in germ cells and erythroid and nonerythroid tissues (6, 14). Approximately 1% of the vertebrate genome is in the HTF fraction, and each island varies in size from about 0.5 to 2.0 kilobases (kb) (5). They are frequently found associated with the 5′ ends of housekeeping genes, e.g., those for hypoxanthine phosphoribosyltransferase (HPRT), adenine phosphoribosyltransferase (APRT), dihydrofolate reductase (DHFR), glucose-6-phosphate dehydrogenase (G6PD), and adenine deaminase (ADA) (5). Indeed, hunting for HTF islands is now used as a method of finding transcriptionally active regions of the genome (13, 28). Highly tissue-specific genes usually lack HTF islands, e.g., β-globin gene family (human, mouse, rabbit, and goat), α-amylase I and II (mouse), growth hormone (mouse), insulin (mouse), and myoglobin (seal). The human α-globin gene is one of a small number of exceptions (6). It seems possible that the transcriptionally promiscuous nature of the human α-globin gene when introduced into heterologous cell lines is associated with the highly GC-rich nature of the promoter. In fact, the 5′-flanking sequence of the human α-globin gene contains six sequences which conform to the consensus SP1-binding site 5′-GGGCGG-3′, known to play an important role in the transcriptional activity of many housekeeping genes (11).

The mouse α-globin gene does not lie in an HTF island, and the promoter region contains no putative SP1-binding sites. It therefore provides an interesting gene system with which to compare the human α-globin gene. Since the transcriptional activity of the mouse α-globin genes has not been studied in heterologous cell lines, we decided to make use of the mouse α-globin gene in an attempt to understand the importance of CpG clusters on transcriptional activity in both heterologous and homologous cell types. We found that in a replicating transfection system, the mouse α-globin gene promoter was more "enhanceable" than that of the human. However, in a nonreplicating system, both the mouse and the human promoter were equally enhancer dependent.

By deletion studies we showed that the 450-base-pair (bp) GC-rich promoter region of the human α-globin gene, which contains all the putative SP1-binding sites, did result in a small (two- to threefold) increase in the transcriptional activity of the promoter in both HeLa cells (nonerythroid) and Putko cells (erythroid). Furthermore, in vitro binding assays suggest that these GC-rich sequences bind SP1 in both cell types. However, even with these 5′ sequences removed, the remaining α-promoter functioned without an SV40 enhancer if the plasmid was replicating. In the absence of replication, the human α-promoter, whether or not it contained the highly GC-rich region, was enhancer dependent just like the mouse α-promoter. It has already been observed that the level of transcriptional activity of the human α-globin promoter is increased manyfold by replication (38). However, our results suggest that this does not reflect a true dependence on active replication, as is found in the case of the adenovirus type 2 (Ad2) major late promoter (MLP) (17) and the Ad2 protein IX promoter (30, 41) but is rather a result of increased copy number in the cell. This result, combined with the observation that in a nonreplicating system the human α-globin promoter does require an SV40 enhancer, suggests that the lack of "enhanceability" in the replicating system is the result of limiting amounts of some *trans*-acting factor, possibly one which binds directly or indirectly to the CCAAT box. Finally, our experiments with the human erythroid cell line Putko suggest that there

* Corresponding author.

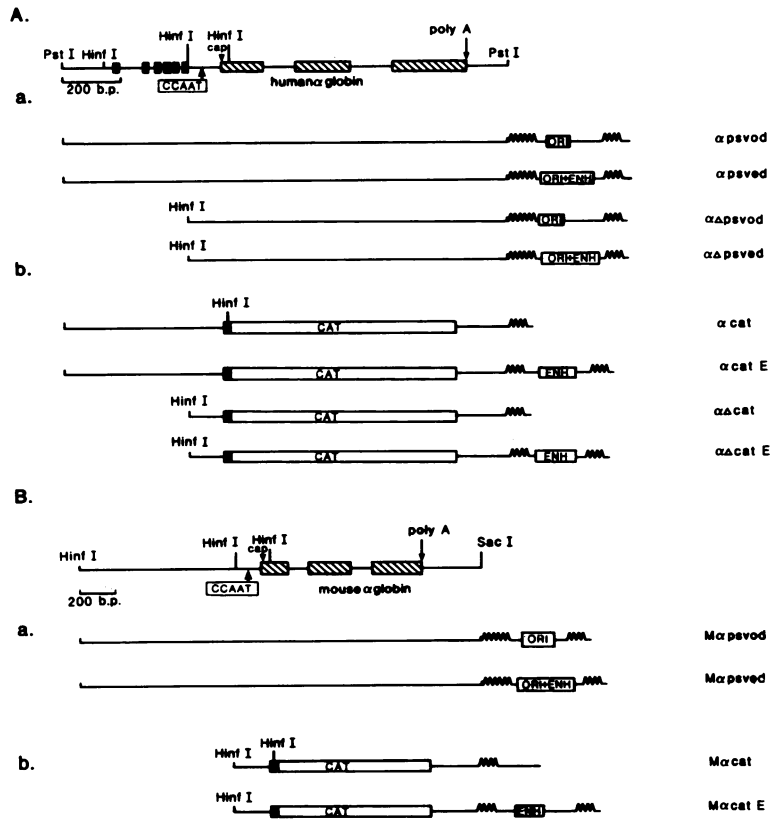

FIG. 1. (A) Line diagram showing some important features of the human α1-globin gene and its flanking DNA. Hatched boxes represent exons; solid boxes represent SP1 consensus sequences. Below (a and b) are maps of all of the plasmids made which contain human α-globin sequences. (B) Line diagram showing some important features of the mouse α-globin gene and its flanking sequences. Hatched boxes represent exons. Below (a and b) are maps of all of the plasmids made which contain mouse α-globin sequences. For more details, see Materials and Methods. poly A, Polyadenylation site; ENH, enhancer.

are no tissue-specific enhancers close to or within the human α-globin gene [from −1500 bp to beyond the poly(A) addition site].

MATERIALS AND METHODS

Constructs. The plasmids αpSVod and αpSVed are made up of the 1.6-kb *Pst* fragment containing the human α1-globin gene (Fig. 1) inserted between the *Pvu*I and *Pst*I sites of the expression vector pSVod (contains the pBR322 tetracycline resistance gene and the SV40 replication origin) (42) and the vector pSVed (identical to pSVod except that the SV40 insert contains both the origin of replication and the SV40 enhancer [35]), respectively.

The plasmids αΔpSVod and αΔpSVed are made up of the α1-globin gene from the *Hin*fI site at −120 bp (cap site is bp 1) to the *Pst*I site at +900 bp cloned in pSVod and pSVed, respectively (see above).

The plasmids αcat and αcatE contain the α1-promoter from the *Hin*fI site at −450 bp to the *Hin*fI site in the 5′ noncoding region cloned into svocat (16) at the *Bgl*II site in the absence (αcat) or presence (αcatE) of the SV40 enhancer cloned into the *Bam*HI site 3′ to the polyadenylation site.

The plasmids αΔcat and αΔcatE contain the *Hin*fI fragment of the α1-globin gene from −110 bp to +20 bp (cap site is bp 1) cloned into the *cat* vectors described above.

The plasmids αcatORI and αΔcatORI are identical to αcat and αΔcat, respectively, except that they contain the SV40 origin of replication cloned into the *Nde*I site of the *cat* vectors.

The plasmids MαpSVod and MαpSVed are made up of the 2-kb fragment containing the mouse α-globin gene (32) from the *Hin*fI site (at −1,000 bp) to the *Sca*I site (at +1,000 bp) cloned into the expression vectors pSVod and pSVed, respectively, at the *Sca*I site.

Plasmids Mαcat and MαcatE contain the *Hin*fI fragment of the mouse α1-globin gene from −140 bp to +10 bp cloned into SV0cat and SV0catE, respectively (see above).

For RβSVpBR328, the rabbit β-globin gene was inserted into the plasmid SVpBR328 containing the SV40 origin and enhancer sequences as well as the large T antigen gene (18).

The plasmid pIRV contains the ampicillin gene, the pBR origin, and the 5′ and 3′ long terminal repeats (LTRs) from Moloney murine leukemia virus. The neomycin resistance gene (*neo*) and the β-galactosidase gene (β-Gal) linked to the rat β-actin promoter were inserted between the LTRs. This plasmid was constructed by J. Morgenstern and H. Land (Imperial Cancer Research Fund, London).

Transient expression. Transfections into HeLa cells were carried out as described previously (35). Plasmid DNA was calcium phosphate precipitated and added to subconfluent dishes of HeLa cells. After 16 h, the medium was changed and the cells were allowed to grow for another 36 h. Transfection into Putko cells was carried out by electroporation (34, 43). Approximately 10^7 cells were harvested in the exponential growth phase, washed with phosphate-buffered saline (PBS), and suspended in 0.8 ml of electroporation buffer (25 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pH 7.1], 140 mM NaCl, 0.75 mM

Na_2HPO_4), and 25 mg of test DNA and 15 μg of pIRV DNA were added. The mixture was electroporated with a BioRad Gene Pulser at 2,000 V and 25 μF, added to 50 ml of medium, and incubated for 48 h at 37°C in 5% CO_2).

RNA purification. The cells were harvested and the RNA was purified as described previously (35). Essentially, cells were lysed in Nonidet P-40 detergent buffer, and the cytoplasmic and nuclear fractions were separated by centrifugation through a sucrose cushion. Following incubation with proteinase K, cytoplasmic RNA was purified by phenol-chloroform extraction and ethanol precipitation.

RNA mapping with S1 nuclease. Probe DNAs (either double or single stranded) were annealed to cytoplasmic RNAs (10 to 20 mg) in 30 μl of 80% formamide–0.04 M PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid), pH 6.4]–0.4 M NaCl–0.1 mM EDTA by denaturing at 80°C for 10 min and then hybridizing at 53°C (double-stranded probe) or 30°C (single-stranded probe) overnight. Then, 300 μl of ice-cold S1 buffer (0.25 M NaCl, 0.03 M sodium acetate [pH 4.6], 2 mM $ZnSO_4$, 50 μg of denatured sonicated carrier DNA per ml, 3,000 U of S1 nuclease) was added and incubated for 1 h at 30°C. S1 reaction mixes were ethanol precipitated and fractionated on denaturing 7 M urea–polyacrylamide gels.

Primer extension. Labeled DNA primer and cold RNA were annealed in 10 μl of 10 mM PIPES (pH 6.4)–0.4 M NaCl at 80°C for 10 min and then hybridized at 63°C overnight. Then, 50 μl of reverse transcriptase buffer (50 mM Tris [pH 8.2], 10 mM dithiothreitol [DTT], 6 mM $MgCl_2$, 0.5 mM each dATP, dCTP, dTTP, and dGTP, 5 U of reverse transcriptase) was added to the hybridization mixes and incubated at 42°C for 1 h. The reaction mixes were ethanol precipitated, RNase treated, and then fractionated by electrophoresis on 7 M urea–polyacrylamide gels.

Assay for CAT. To assay for chloramphenicol acetyltransferase (CAT), cells were harvested for assay 48 h after transfection and washed in PBS. The pelleted cells were sonicated in 100 μl of 0.25 M Tris hydrochloride (pH 7.8) and spun in an Eppendorf microfuge for 10 min, and the supernatant was removed and assayed for enzyme activity with [^{14}C]chloramphenicol and unlabeled acetyl coenzyme A (16).

Assay for β-galactosidase activity. The β-galactoside activity used to monitor transfection efficiency for each dish was assayed according to Herbomel et al. (20). A 1-ml amount of 60 mM Na_2HPO_4–40 mM NaH_2PO_4–10 mM KCl–1 mM $MgCl_2$–50 mM β-mercaptoethanol–0.2 ml of *o*-nitrophenyl-β-D-galactopyranoside was added to 50 μl of cell extract. The reaction was carried out at 37°C for 1 h and stopped by adding 0.5 ml of 1 M $NaCO_3$. The OD_{420} was then measured.

Preparation of nuclear extract. Nuclear extracts from HeLa cells were prepared as described previously by Dignam et al. (10). The final protein concentration was 3 to 4 mg/ml.

Electrophoretic mobility shift binding assay. DNA inserts were isolated from agarose gels, and the ends were filled in with [^{32}P]dATPs by using the large fragment of DNA polymerase I. Then, 1 ng of ^{32}P-labeled DNA, 1 μg of poly(dI-dC), and 1 μl of nuclear extract (~3 μg of protein) were mixed together in a final volume of 25 μl of factor-binding buffer (10 mM Tris[pH 7.4], 1 mM EDTA, 5% glycerol, 1 mM DTT). The final concentration of NaCl was varied from 50 to 120 mM. After incubation for 20 min at 30°C, the reaction mixes were loaded directly onto low-ionic-strength (0.25× TBE) 4% polyacrylamide gels and run at 15 V/cm for 3 to 5 h. Gels were dried and subjected to autoradiography overnight.

RESULTS

Expression of mouse α-globin gene in a heterologous cell line. Since the mouse α-globin gene lacks the CpG cluster found in the promoter of the human α-globin gene, we decided to investigate the transcriptional activity of this gene in a heterologous cell type (HeLa cells) and to determine its requirement for an SV40 enhancer. Initially we constructed two clones, MαpSVod and MαpSVed, which consist of a 2.5-kb fragment containing the entire mouse α-globin gene with approximately 1 kb of 5′-flanking sequence cloned into the expression vectors pSVod and pSVed, respectively (Fig. 1). The plasmid pSVed contains the SV40 enhancer and origin of replication, while pSVod contains only the SV40 origin of replication (35). HeLa cells were transfected with MαpSVod and MαpSVed along with the plasmid RβSV pBR328 (see Materials and Methods), which acts both as a source of large T antigen and as a cotransfection control. The plasmid RβSVpBR328 contains the rabbit β-globin gene and the SV40 large T antigen gene (see Materials and Methods). The presence of large T antigen results in the replication of any plasmid containing the SV40 origin (e.g., MαpSVod and MαpSVed). The cells were harvested 48 h later, and the RNA was analyzed for levels of mouse α-globin at both the 5′ end (by primer extension) and the 3′ end (by S1 nuclease analysis) (Fig. 2). The probes used are shown in the lower half of Fig. 2. Mouse α-globin mRNA started at the predicted cap site (Fig. 2A) and ended at the predicted polyadenylation site (Fig. 2B). The level of α-globin-specific mRNA after transfection with the SV40 enhancer-containing construct (MαpSVed) was at least 50-fold greater than that found after transfection with the construct without the enhancer (MαpSVod). The level of expression of the rabbit β-globin gene was equal in both cases, showing that the transfections were equally efficient. This level of response to the SV40 enhancer was much greater than that found by us (Fig. 3) and by others (38) for the human α-globin gene. In the absence of replication, the level of mouse α-globin mRNA was too low to be analyzed accurately.

Effect of upstream GC-rich sequences on human α-globin gene expression in a heterologous cell line. In an attempt to discover whether the highly GC-rich nature of the human α-globin promoter was associated with the high rates of transcription in heterologous cells in the absence of an enhancer, we constructed the plasmids αΔpSVod and αΔpSVed, in which the promoter region was reduced to a 130-bp fragment lacking most of the GC-rich sequence and all of the putative SP1-binding sites. The transcriptional activities of these constructs were compared with that of αpSVod and αpSVed, in which the entire 600 bp of 5′-flanking sequence was present (Fig. 1). These constructs were transfected into HeLa cells in the presence of RβSVpBR328 (see Materials and Methods) to allow replication of the SV40 *ori*-containing plasmids. Cytoplasmic RNA was prepared 48 h later and analyzed for α-globin mRNA at both the 5′ end (by primer extension) and the 3′ end (by S1 nuclease). The results are shown in Fig. 3. First, it is clear that the α-globin gene worked well without any enhancer, and if the level of mRNA produced from αpSVod is compared with that from αpSVed, the presence of the SV40 enhancer increased transcription no more than three- to fourfold. This agrees with the findings of others (38). Furthermore, the level of α-globin mRNA was reduced approximately threefold by the deletion of the 450-bp CpG-rich region (compare αpSVod with αΔpSVod and αpSVed with

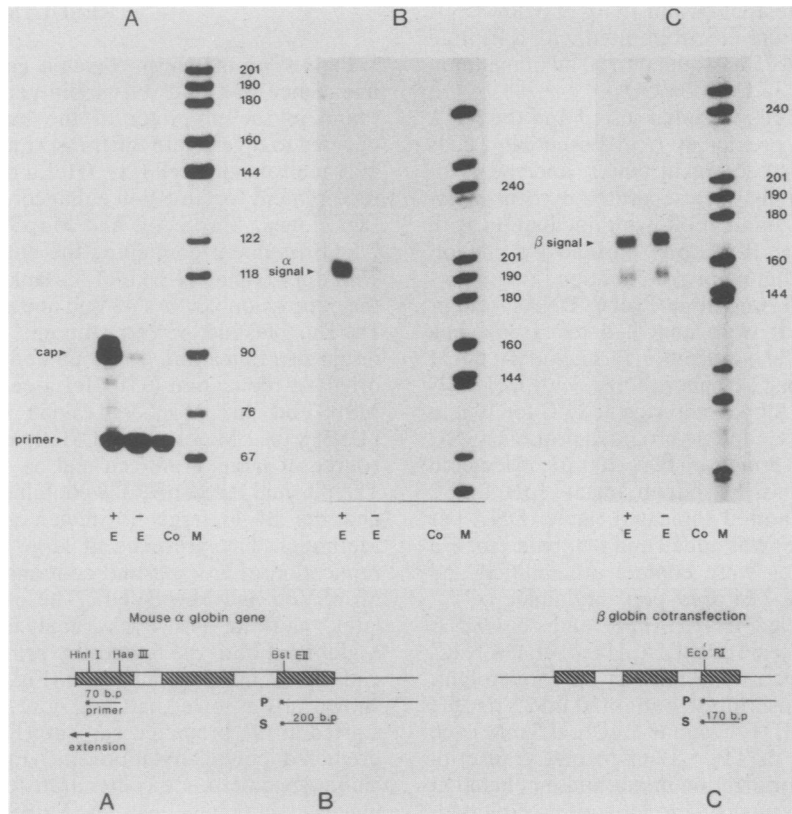

FIG. 2. (A) 5′-End primer extension analysis of cytoplasmic mRNA from HeLa cells transfected with the mouse α-globin gene via either MαpSVod (−E) or MαpSVed (+E). Co, Minus-RNA control. (B) 3′-End S1 nuclease analysis of cytoplasmic mRNA from HeLa cells transfected with the mouse α-globin gene. (C) 3′-End analysis of cytoplasmic mRNA from same transfection as in A and B but tested for the level of rabbit β-globin mRNA (cotransfection control). In the line diagrams below, S denotes signal, P denotes probe, and the asterisk denotes the position of the label.

αΔpSVed). However, it is also clear that this deletion did not convert the human α-globin promoter into an enhancer-dependent promoter such as the mouse α-promoter (Fig. 2).

Transcriptional activity of the human and mouse α-globin gene promoters in a nonreplicating system. In order to verify that the decrease in the level of α-globin mRNA from the deleted promoter was a true effect of the level of transcription and was not related to a possible decrease in the copy number of the plasmid, we decided to assay for transcriptional activity in the absence of replication. To obtain signals that were high enough to be detected, we decided to use the CAT assay system developed by Gorman et al. (16). The *cat* gene is placed under the control of the promoter in question in either the presence or absence of the SV40 enhancer, and the construct is transfected into cells in culture (see Materials and Methods). After 48 h, the enzymatic activity is measured in cell extracts by estimation of the conversion of [^{14}C]chloramphenicol into its acetylated form. This value is taken as a measure of the transcriptional activity of the promoter.

In the first series of experiments, we measured the relative transcriptional activity of the deleted human α-globin promoter in the absence (αΔcat) and presence (αΔcatE) of the SV40 enhancer and compared this with the activity found for the larger α-globin promoter fragment, again in the absence (αcat) and presence (αcatE) of the enhancer. A similar construct was made with the mouse α-globin promoter by placing a 150-bp *Hin*fI fragment in front of the *cat* gene (Mαcat). For details of these constructs, see Fig. 1. The plasmid SV2cat, which contains both the SV40 early promoter and enhancer, was used as the positive control (see Materials and Methods). This promoter is known to function efficiently in many cell lines (16). The vectors with and without the enhancer (svocatE and svocat, respectively) but lacking any promoter element were used as the negative controls. Variations in cell density were controlled by assaying the cell lysates for protein content, and variations in transfection efficiency were controlled by cotransfection with the β-actin/β-galactosidase plasmid pIRV (see Materials and Methods).

Figure 4A shows a representative example of the results obtained. The experiments were repeated three times and always gave the same pattern. Transcriptional activity was not observed with any promoter fragment in the absence of the SV40 enhancer. The activity was increased at least 20-fold by the SV40 enhancer. Furthermore, the level of transcriptional activity with the large human α-promoter fragment (αcatE) was consistently two to three times that found with the deleted promoter (αΔcatE), as we have found for the intact α-globin gene (Fig. 3). There are two points to be made here. First, in the absence of replication, the human α-globin promoter was enhanced greatly by the SV40 promoter. In fact, it behaved in a similar way to the mouse α-globin promoter (Mαcat and MαcatE). Second, the 450-bp GC-rich region appeared to have only a slight (two- to threefold) positive effect on transcription.

It remained theoretically possible that the reason for the high level of expression of the human α-globin gene after

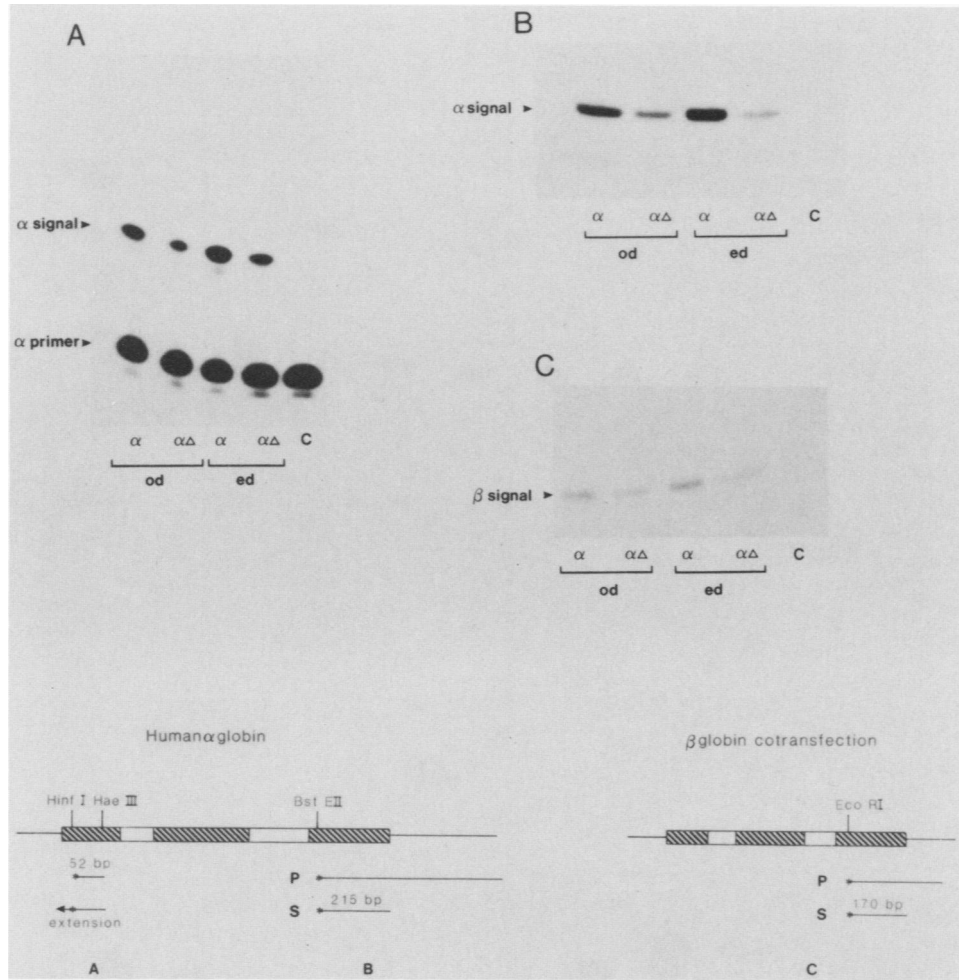

FIG. 3. (A) 5′-End primer extension analysis of cytoplasmic mRNA from HeLa cells transfected with the human α-globin gene via αpSVod (α.od), αΔpSVod (αΔ.od), αpSVed (α.ed), and αΔpSVed (αΔ.ed). C denotes minus-RNA control. (B) 3′-End S1 analysis of the cytoplasmic RNA described above. (C) 3′-End S1 analysis of the cytoplasmic RNA described above, with rabbit β-globin probe (cotransfection control). In the line diagrams below, S denotes signal, P denotes probe, and the asterisk denotes the position of the label. Hatched boxes represent exons.

replication in the absence of an SV40 enhancer (as shown in Fig. 3) was due to some sequence within the body of the gene which could act as an enhancer in HeLa cells. The experiments with the CAT system did not rule this out, since the plasmids used did not contain the α-globin coding sequences. To address this question, we used the enhancer-trap plasmid SV1cat (16). This plasmid contains the SV40 promoter without the SV40 enhancer, and transcription from this promoter is very low. By inserting DNA sequences of interest in front of this promoter, one can test whether they can act as enhancers. Insertion of the entire body of the α-globin gene from the initiation site (*Nco*I) to beyond the polyadenylation site (*Pst*I) in either orientation (SV1catNP and SV1catPN) (see Materials and Methods) had no effect on the transcriptional activity of SV1cat (data not shown), suggesting the absence of any enhancerlike sequences.

Effect of replication on transcriptional activity of the human α-globin promoter. Using the CAT system, we were keen to discover whether αcat and αΔcat (plasmids lacking the SV40 enhancer) would become transcriptionally active in the presence of replication. For this reason, we constructed the plasmids αcatORI and αΔcatORI, which have an intact SV40 origin of replication but lack the SV40 enhancer (see Materials and Methods). When these constructs were cotransfected with RβSVpBR328, they both gave equally high levels of expression (data not shown). The presence of the SV40 *ori* fragment in the negative control plasmid SV0cat (SV0catORI) (see Materials and Methods) did result in higher background if cotransfected with the T antigen-containing plasmid, but only 5% of that observed when the α-globin promoter was present (Fig. 4B). We carried out a time course trial in order to discover whether this increase in CAT activity was simply the result of increased copy number or whether it was the result of an increased rate of transcription off each plasmid due, for example, to some conformational change in the DNA following replication. The plasmids αcatORI and svocatORI were transfected into HeLa cells in the presence of RβSVpBR328, and cell lysates were made at 6, 12, 24, and 48 h after transfection. The results are shown in Fig. 4B. It is clear that the major increase in CAT activity did not occur until the 48-h time point. Since plasmids transfected into HeLa cell nuclei started to replicate well before this time, we would expect the CAT activity increase to occur earlier if it was associated with a conformation change in the DNA resulting from a single round of replication. Instead, the increase in CAT

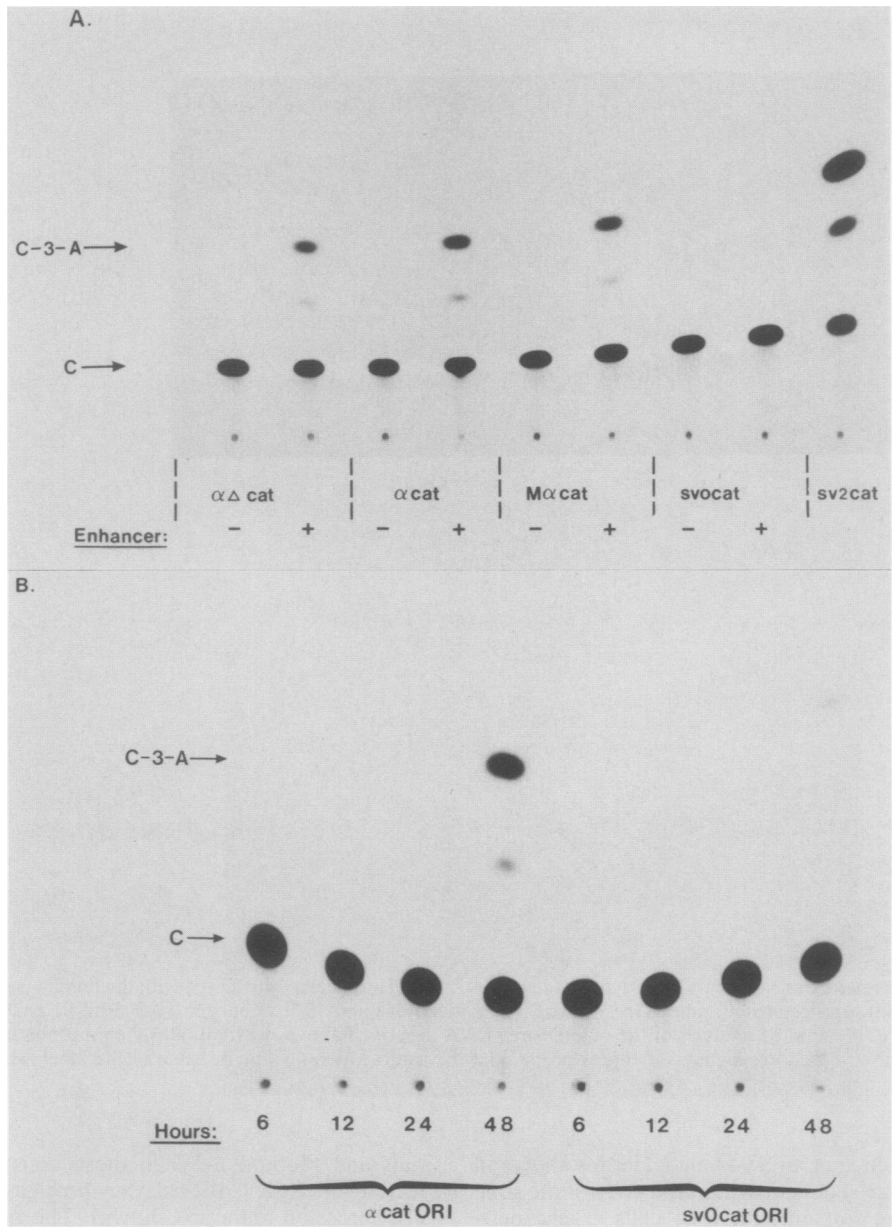

FIG. 4. (A) CAT assay of cell lysates from HeLa cells transfected with αΔcat, αcat, Mαcat, and SV0cat in either the absence (−) or presence (+) of the SV40 enhancer. SV2cat is the positive control for the transfection and CAT assay. C denotes [^{14}C]chloramphenicol, and C-3-A denotes its acetylated product. (B) CAT assays of cell lysates from HeLa cells transfected with αcatORI and SV0catORI. Both were cotransfected with RβSVpBR328 to provide large T antigen for replication. Lysates were prepared 6, 12, 24, or 48 h after transfection.

activity appeared to be exponential, mimicking the pattern found with the negative control svocatORI, which contains no specific promoter element. From these experiments it seems highly likely that the human α-globin promoter is not replication dependent in the manner observed for the Ad2 MLP (17) and the Ad2 protein IX promoter (30, 41). This is consistent with the fact that the human α-globin promoter operated in the absence of replication provided that the SV40 enhancer was present (Fig. 4).

Transcriptional activity of the human α-globin promoter in an erythroid cell line. It was important to test the behavior of the various human α-globin subclones in an erythroid cell line in an attempt to find regions of DNA which may be involved in tissue-specific regulation. Using the technique of electroporation, we found that we could successfully transfect Putko cells (a derivative of K562 cells known to produce significant amounts of human α-globin mRNA) (24, 26) and could obtain reproducible results with the CAT vectors (αΔcat, αcat, and Mαcat). No activity was seen in the absence of an SV40 enhancer. The results of one representative experiment are shown in Fig. 5. The levels of CAT activity in the presence of the enhancer were consistently higher (two- to threefold) when the α-globin promoter included the GC-rich upstream region. This result therefore mimics those found in nonerythroid HeLa cells. It is interesting that the ratio of SV2cat activity to αcatE activity was 10-fold lower in Putko cells (~0.3) than in HeLa cells (~3.0), but we do not know whether this reflects a greater level of activity of the α-globin promoter in erythroid cells or a decreased activity of the SV40 early promoter in these cells.

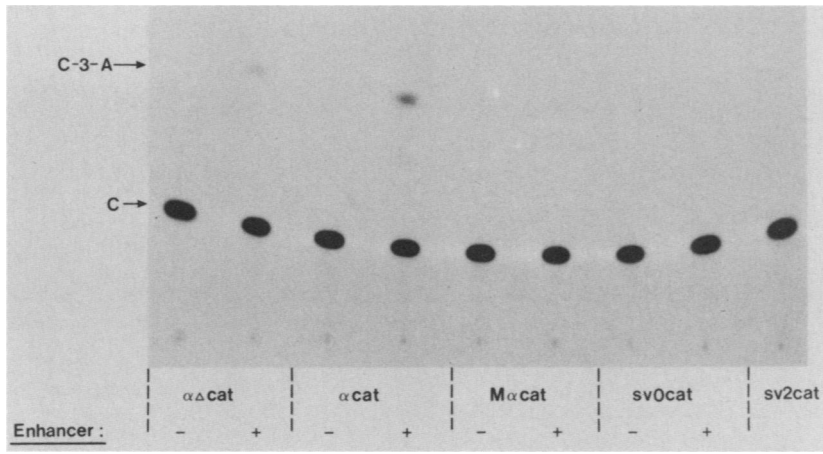

FIG. 5. CAT assay of cell lysates from Putko cells transfected with αΔcat, αcat, Mαcat, and SV0cat in either the absence (−) or presence (+) of the SV40 enhancer. C denotes [^{14}C]chloramphenicol, and C-3-A denotes its acetylated product.

The fact that the Putko cell line produces endogenous human α-globin (data not shown) but that the αcat construct is not transcriptionally active in the absence of an SV40 enhancer suggests that the 600-bp fragment from *Pst*I in the 5′-flanking region to the *Nco*I site is not sufficient to confer the tissue-specific regulation of this gene. Other globin genes, when transfected into K562 or Putko cells, behave like their endogenous counterparts, e.g., the β-globin gene or promoter region is not transcriptionally active, while ε-globin and ζ-globin are active (1, 24, 25, 37). Experiments carried out in this laboratory on the ζ-globin promoter also show tissue-specific expression in Putko cells (P. Lamb, personal communication).

In an attempt to look for tissue-specific enhancers within the body of the α-globin gene, we used the plasmids SV1catNP and SV1catPN, containing DNA from the initiation codon (*Nco*I site) to beyond the polyadenylation site (*Pst*I) in both orientations (as described in Materials and Methods). No increase in activity was observed (data not shown). We also constructed two more plasmids to test for tissue-specific enhancers further 5′ to the *Pst*I site at −600 bp. These constructs, sv1catBS and sv1catSB, contained the 1-kb fragment from *Bgl*II at −1400 to *Sma*I at −670 in both orientations (see Materials and Methods). Again, no increase in the activity of the SV40 promoter was observed after transfection into Putko cells (data not shown). These results suggest that the entire region from 2 kb 5′ of the cap site to beyond the polyadenylation site does not contain any strong tissue-specific enhancers.

In vitro analysis of sequence-specific DNA-binding proteins. Bird (5) has put forward a model to explain the existence of HTF islands. He suggests that the maintenance of hypomethylated HTF islands is achieved by bound factors (proteins) which sterically exclude the methylase. These factors presumably have an affinity for sequences containing CpG. This fact, combined with the knowledge that the region of the human α-globin promoter which contained sequences which should bind SP1 has a positive effect on transcription, meant that it was of interest to study the human α-promoter region at the level of sequence-specific DNA-binding proteins. An electrophoretic mobility shift assay (15) was used to resolve specific protein-DNA interactions. In this assay, ^{32}P-labeled DNA fragments were incubated with nuclear extracts (10) from HeLa cells in the presence of various amounts of unlabeled competitor DNA. After electrophoresis on low-ionic-strength polyacrylamide gels and autoradiography, stable protein-DNA complexes that were formed during the incubation were visualized as discrete bands with reduced mobility compared with free DNA.

The results of these experiments are shown in Fig. 6. When the 130-bp *Hin*fI fragment containing the CCAAT and ATA boxes was incubated with HeLa cell nuclear extract, it formed two complexes (A and B). The presence of these two bands does not necessarily imply the binding of two different factors; one band could be a degradation product of the other. A preliminary salt curve showed that these complexes were stable at salt concentrations from 50 to 125 mM and that their presence was inhibited by excess cold *Hin*fI fragment (data not shown).

In order to define precisely which region of this fragment was involved in the binding, assays were repeated in the presence of unlabeled oligonucleotides to putative regions of importance. Oligonucleotide I is a 30-mer covering the sequence from −36 to −66 (Fig. 6C). This sequence is highly GC rich and contains the sequence GCCGCCCG, which is very similar but not identical to the opposite strand of the consensus SP1-binding site (G/T)GGGCGG(G/A) [(T/C)CC GCCC(C/A)]. Furthermore, this sequence appears to be unique to the human α-globin promoter among globin gene promoters. It is not found in an equivalent position in the mouse α-globin gene promoter or in the human β-globin gene promoter. Oligonucleotide II is a 15-mer containing the CCAAT box, which is known to be involved in the binding of proteins which are important in transcriptional activity in many eucaryotic promoters (22), and deletion studies of the human α-globin gene promoter have shown it to be functionally important in HeLa cells (31). As shown in Fig. 6A, oligonucleotide I did not compete with the specific binding, while oligonucleotide II competed with both stable complexes A and B at levels as low as 50 ng. This suggests that the highly GC-rich sequence which lies between the CCAAT and ATA boxes is not involved in factor binding. This result is consistent with the observation (Fig. 4) that the small mouse α-globin promoter plasmid (Mαcat), which does not possess the GC-rich element, and the small human α-globin promoter plasmid (αΔcat) behaved in a similar manner.

It was also of interest to study the 300-bp upstream fragment, which we have shown had a small but reproducible effect on transcriptional activity, and more specifically to ask whether the SP1 protein binds to any or all of the

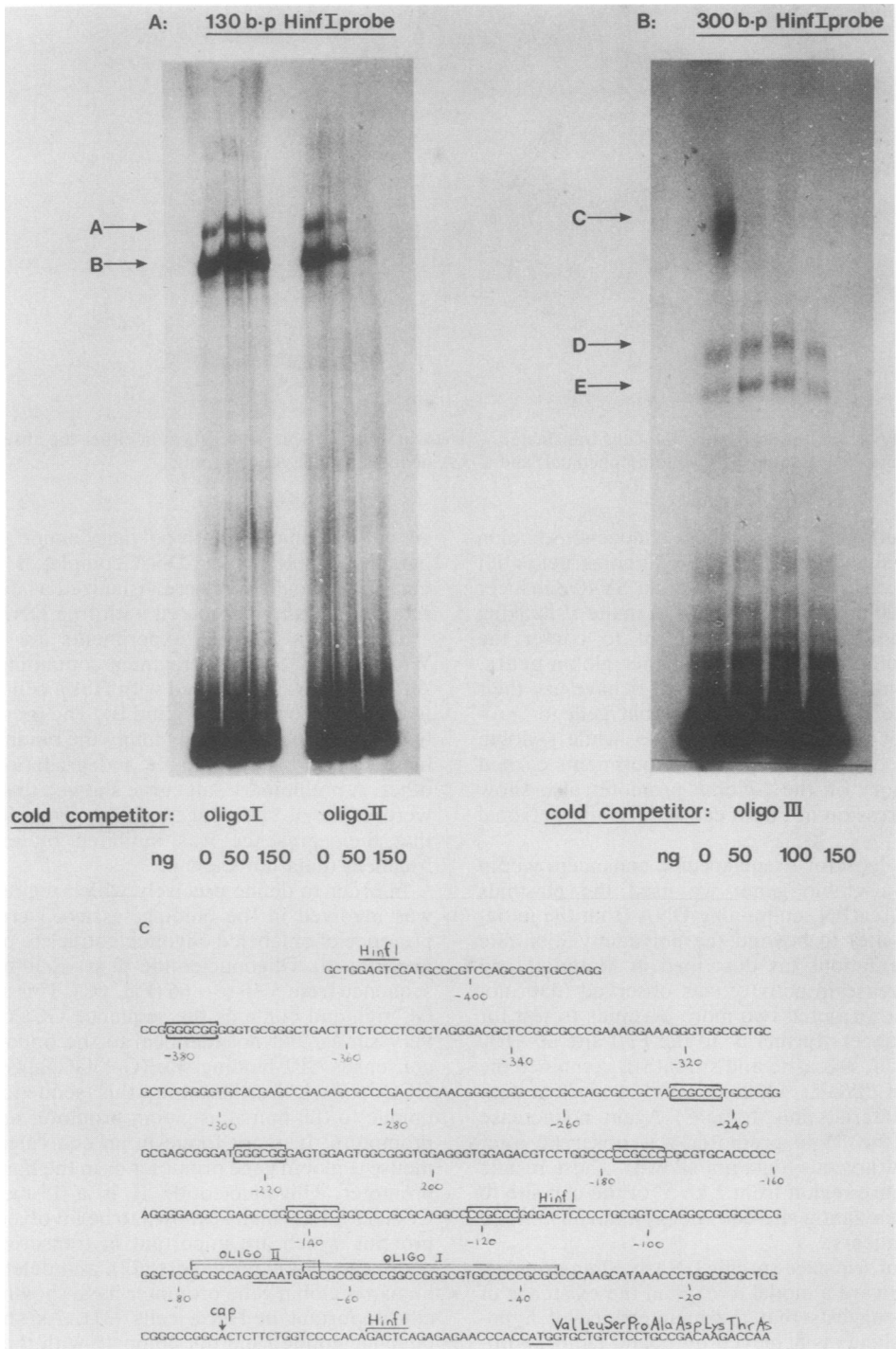

FIG. 6. Gel retardation assay carried out on the 130-bp *Hin*fI fragment in the presence of cold competitor (oligonucleotide I or II). Bands A and B are indicated. (B) Gel retardation assay carried out on the 300-bp *Hin*fI fragment in the presence of cold competitor oligonucleotide III (5'-GGAACTGGGCGGAGTTAGG-3', see text). Bands C, D, and E are indicated. (C) Sequence of the promoter region of the human α1-globin gene. Cap denotes the start of transcription; the translational start site ATG is underlined. The sequences containing the core SP1-binding sequences are boxed.

putative SP1-binding sites. The 300-bp *Hin*fI fragment (Fig. 6C) did form stable complexes with factors in the nuclear extract (C, D, and E in Fig. 6B). These complexes were stable at 50 to 125 mM NaCl and could be competed with by excess cold *Hin*fI fragment (data not shown). Oligonucleotide III was a 19-mer (5'-GGAACTGGGCGGAGTTAGG-3') identical to the sequence in the SV40 21-bp repeat which has been shown to bind SP1 in vitro with a high affinity (23). This oligonucleotide competed with major band C even at levels as low as 50 ng. This suggests that SP1 does bind to sequences within the 300-bp fragment, presumably to one, some, or all six of the regions containing the (G/T)GGGCGG

(G/A) sequence (Fig. 5C). Further analysis is needed to define precisely which sequences are involved. It is also of interest that bands D and E were not competed with by oligonucleotide III, suggesting that other sequences within the 300-bp region are forming specific stable complexes with proteins. It would obviously be of interest to establish whether these represent specific HTF island DNA-binding proteins, as proposed by Bird (5).

DISCUSSION

These studies were principally instigated to investigate the marked transcriptional promiscuity of the human α-globin gene. First, we show that the mouse α-globin gene behaves differently from the human α-globin gene. In the presence of 1 kb of 5'-flanking sequence, the presence of the SV40 enhancer increased the rate of transcription from the mouse α-globin promoter 50-fold but increased the rate of transcription from the human α-promoter only 3- to 4-fold. However, this difference was only observed when the plasmids were replicating in the cell. The possibility that the presence of a hypomethylated HTF island at the 5' end of the human but not the mouse α-globin gene accounted for this difference was then investigated. Deletion studies revealed that although the upstream GC-rich region (HTF island) of the α1-globin promoter did have a positive effect on transcription from the α-globin promoter, this increase was small (two- to threefold). In vitro binding studies revealed that this region bound SP1 and other unknown DNA-binding proteins from HeLa cell extracts, suggesting that these DNA-binding proteins may be associated with the role of this region in increasing the rate of transcription. However, it is also clear that the GC-rich region was not responsible for the promiscuous nature of the human α-globin gene promoter in a replicating system, since deletion of this region resulted in a promoter which still functioned well in the absence of an SV40 enhancer provided that we assayed for this in a replicating system. By utilizing the CAT system, in which increased sensitivity allowed us to use a nonreplicating system, we found that the human α-promoter, both with and without the upstream flanking sequences, was enhanceable (ca. 20-fold) and was just as enhanceable as the mouse α-promoter.

It therefore follows that the human α-globin promoter, like that of the mouse, does require an SV40 enhancer to function in heterologous cells unless the plasmid is replicating. The idea that replication per se may activate the behavior of the human α-globin promoter is at first sight attractive, since it is known that the time in differentiation at which globin gene expression is switched on during erythropoiesis (the proerythroblast) is a time of rapid cell division (29). DNA viruses display dramatically different patterns of gene expression before and after replication, and recent studies on the Ad2 MLP (17), the pIX promoter (30, 41), and the *Xenopus laevis* β-globin gene (12) suggest that active replication is required for these promoters to be active. However, the time course experiments presented here (Fig. 4B) suggest that the increase in transcription from the α-promoter which is associated with replication is simply the result of increased copy number. Furthermore, it is worth remembering that transcriptional activity from the human β-globin promoter is not increased by replication (38). If replication did play an important role in the control of globin gene expression during the development of the proerythroblast, one might expect that it would act on all globin genes alike.

One possible explanation for the observed activation of the human α-globin gene at high copy number is that *trans*-acting factors regulating its transcription bind to the human α-promoter only weakly in the absence of an enhancer. However, at high copy number, the cumulative effect of these weak interactions is sufficient to allow significant levels of enhancer-independent transcription. It is possible that the CCAAT box-binding factor detected in the gel retardation experiments is responsible for this effect. It is now well established that the CCAAT promoter sequence is functionally involved in the transcription of many genes and that proteins binding to it may therefore be present in excess in the cell. Indeed, it is now clear that there is not just one CCAAT box-binding protein but a family of them, each binding to only a subset of the genes containing the CCAAT box sequence (36). At present we can only speculate as to which of these factors confers this unusual transcriptional behavior on the human α-globin gene.

Could the behavior of the human α-promoter in nonerythroid cells have any bearing on its behavior in vivo, where it is regulated in a tissue-specific manner? To address this issue, we have studied the behavior of human α-promoter-containing constructs in an erythroid cell line (Putko) and found that the transcriptional activity mimicked that in the nonerythroid cell line (HeLa). Observable transcription required either an SV40 enhancer (Fig. 5) or replication (data not shown), and the upstream GC-rich region increased transcriptional activity two- to threefold (Fig. 5). Furthermore, no erythroid-specific enhancers could be found in or around the gene. This suggests that any erythroid-specific elements lie some distance from the gene itself. Erythroid-specific expression of the human β-globin gene has been shown both in mouse erythroleukemia (MEL) cells and in transgeneic mice to involve sequences at both the 5' end (between −180 and −150) and the 3' end (ca. 0.5 to 1.2 kb beyond the polyadenylation site) (2, 4, 7, 27, 44). Comparison of the region between −180 and −150 of the human β-globin gene with this region in other β-globin genes reveals the presence of a conserved sequence at −160, 5'-(A/G)A (C/T) . . . TC(T/C)TAAGCCAGTGCCA-3'.

This sequence is not found in the 5'-flanking sequence of the human α-promoter, and this may explain the absence of any tissue-specific enhancer in this region. However, recent experiments have defined an additional very strong erythroid-specific enhancer which lies 50 kb 5' and 20 kb 3' to the β-globin gene at the site of a cluster of erythroid-specific DNase I-hypersensitive sites (19, 39). This region, originally shown to enhance the expression of the *cat* gene in K562 cells (40), has recently been shown to act as a tissue-specific enhancer in transgeneic mice (19). The presence of these regions increases the transcriptional activity of the human β-globin gene in transgeneic mice by at least 20-fold, and their action is tissue specific. A similar study of the chromatin of the α-locus revealed strong DNase I-hypersensitive sites at −300 to −100 bp to the α-globin gene's cap site (45), which was erythroid specific. However, this region of DNA failed to enhance the expression of SV1cat in either HeLa cells or Putko cells (see Results). The sequences responsible for the tissue-specific expression of the α-globin genes remain undetected and may lie a long distance from the genes themselves. This would fit with the fact that a 14-kb fragment containing both of the human α-globin genes is not functionally active in the erythroid cells of transgeneic mice (33). Experiments with MEL cells, in which induction with dimethyl sulfoxide results in erythroid differentiation (involving a 50- to 100-fold increase in the expression of the

endogenous globin genes), have found that the expression of exogenous β-globin genes mimics this induction (44). However, the only experiment in which human α-globin expression has been shown to be inducible in MEL cells is one in which the entire chromosome 16 was present (9), whereas when small 5-kb fragments containing the α1-globin gene were used, no inducibility was shown (7).

ACKNOWLEDGMENTS

This work was supported by a project grant (G870893CB) from the Medical Research Council to E.W. and N.J.P.

We thank P. Lamb for helpful discussions and H. Land (Imperial Cancer Research Fund) for the pIRV construct.

LITERATURE CITED

1. **Allan, M., P. Montague, G. J. Grindlay, G. Sibbet, M. Donovan-Peluso, A. Bank, and J. Paul.** 1985. Tissue specific transcription of the human ε-globin gene following transfection into the embryonic erythroid cell line K562. Nucleic Acids Res. **13:** 6125–6136.
2. **Antoniou, M., E. deBoer, G. Habets, and F. Grosveld.** 1988. The human β-globin gene contains multiple regulatory regions: identification of one promoter and two downstream enhancers. EMBO J. **7:**377–384.
3. **Banerji, J., S. Rusconi, and W. Schaffner.** 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell **27:**299–308.
4. **Behringer, R. R., R. E. Hammer, R. L. Brinster, R. D. Palmiter, and T. M. Townes.** 1987. Two 3′ sequences direct adult erythroid-specific expression of human β-globin genes in transgenic mice. Proc. Natl. Acad. Sci. USA **84:**7056–7060.
5. **Bird, A. P.** 1986. CpG-rich islands and the function of DNA methylation. Nature (London) **321:**209–213.
6. **Bird, A. P., M. H. Taggart, R. D. Nicholls, and D. R. Higgs.** 1987. Non-methylated CpG-rich islands at the human α-globin locus: implications for evolution of the α-globin pseudogene. EMBO J. **6:**999–1004.
7. **Charnay, P., R. Treisman, P. Mellon, M. Chao, R. Axel, and T. Maniatis.** 1984. Differences in human α- and β-globin gene expression in mouse erythroleukemia cells: the role of intragenic sequences. Cell **38:**251–263.
8. **Cooper, D. N., M. H. Taggart, and A. P. Bird.** 1983. Unmethylated domains in vertebrate DNA. Nucleic Acids Res. **11:**647–658.
9. **Deisseroth, A., and D. Hendrick.** 1978. Human α globin gene expression following chromosomal dependent gene transfer into mouse erythroleukemia cells. Cell **15:**55–63.
10. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475–1489.
11. **Dynan, W. S., and R. Tjian.** 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature (London) **316:**774–777.
12. **Enver, T., A. C. Brewer, and R. K. Patient.** 1988. Role for DNA replication in β-globin gene activation. Mol. Cell. Biol. **8:**1301–1308.
13. **Estivill, X., M. Farrall, P. J. Scambler, G. M. Bell, K. M. F. Hawley, N. J. Lench, G. P. Bates, H. C. Kruyer, P. A. Frederick, P. Stainer, E. K. Watson, R. Williamson, and B. J. Wainwright.** 1987. A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. Nature (London) **326:** 840–845.
14. **Fischel-Ghodsian, N., R. D. Nicholls, and D. R. Higgs.** 1987. Unusual features of CpG-rich (HTF) islands in the human α globin complex: association with nonfunctional pseudogenes and presence within the 3′ portion of the ζ gene. Nucleic Acids Res. **15:**9215–9225.
15. **Fried, M. G., and D. M. Crothers.** 1984. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. J. Mol. Biol. **172:**263–282.
16. **Gorman, C. M., L. Moffat, and B. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
17. **Grass, D. S., D. Read, E. D. Lewis, and J. L. Manley.** 1987. Cell- and promoter-specific activation of transcription by DNA replication. Genes Dev. **1:**1065–1074.
18. **Grosveld, G. C., E. de Boer, C. K. Shewmaker, and R. A. Flavell.** 1982. DNA sequences necessary for transcription of the rabbit β-globin gene in vivo. Nature (London) **295:**120–126.
19. **Grosveld, F., G. Blom van Assendeft, D. R. Greaves, and G. Kollias.** 1987. Position-independent, high-level expression of the human β-globin gene in transgenic mice. Cell **51:**975–985.
20. **Herbomel, P., B. Bourachot, and M. Yaniv.** 1984. Two distinct enhancers with different cell specificities coexist in the regulatory region of polyoma. Cell **39:**653–662.
21. **Humphries, R. K., T. Ley, P. Turner, A. D. Moulton, and A. W. Nienhuis.** 1982. Differences in human α-, β- and δ-globin gene expression in monkey kidney cells. Cell **30:**173–183.
22. **Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian.** 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell **48:**79–89.
23. **Kadonaga, J. T., K. A. Jones, and R. Tjian.** 1986. Promoter-specific activation of RNA polymerase II transcription by Sp1. TIBS **11:**20–23.
24. **Khazaie, K., F. Gounari, M. Antoniou, E. deBoer, and F. Grosveld.** 1986. β-Globin promoter generates 5′-truncated transcripts in the embryonic/fetal erythroid environment. Nucleic Acids Res. **14:**7199–7212.
25. **Kioussis, D., F. Wilson, K. Khazaie, and F. Grosveld.** 1985. Differential expression of human globin genes introduced in K562 cells. EMBO J. **4:**927–931.
26. **Klein, G., J. Zeuthen, I. Eriksson, P. Terasaki, M. Bernoco, A. Rosen, G. Masucci, S. Povey, and R. Ber.** 1980. Hybridization of a myeloid leukemia-derived human cell line (K562) with a human Burkitts lymphoma line (P3HR-1). J. Natl. Cancer Inst. **64:**725–738.
27. **Kollias, G., N. Wrighton, J. Hurst, and F. Grosveld.** 1986. Regulated expression of human $^{A}\gamma$-, β-, and hybrid γβ-globin genes in transgenic mice: manipulation of the developmental expression patterns. Cell **46:**89–94.
28. **Lavia, P., D. Macleod, and A. Bird.** 1987. Coincident start sites for divergent transcripts at a randomly selected CpG-rich island of mouse. EMBO J. **6:**2773–2779.
29. **Marks, P. A., and R. A. Rifkind.** 1978. Erythroleukemic differentiation. Annu. Rev. Biochem. **47:**419–448.
30. **Matsui, T., M. Murayama, and T. Mita.** 1986. Adenovirus 2 peptide IX gene is expressed only on replicated DNA molecules. Mol. Cell. Biol. **6:**4149–4154.
31. **Mellon, P., V. Parker, Y. Gluzman, and T. Maniatis.** 1981. Identification of DNA sequences required for transcription of the human α1-globin gene in a new SV40 host-vector system. Cell **27:**279–288.
32. **Nishioka, Y., and P. Leder.** 1979. The complete sequence of a chromosomal mouse α globin gene reveals elements conserved throughout vertebrate evolution. Cell **18:**875–882.
33. **Palmiter, R. D., and R. L. Brinster.** 1986. Germ-line transformation of mice. Annu. Rev. Genet. **20:**465–499.
34. **Potter, H., L. Weir, and P. Leder.** 1984. Enhancer-dependent expression of human K immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. Proc. Natl. Acad. Sci. USA **81:**7161–7165.
35. **Proudfoot, N. J., T. R. Rutherford, and G. A. Partington.** 1984. Transcriptional analysis of human ζ globin genes. EMBO J. **3:** 1533–1540.
36. **Raymondjean, M., S. Cereghini, and M. Yaniv.** 1988. Several distinct "CCAAT" box binding proteins coexist in eukaryotic cells. Proc. Natl. Acad. Sci. USA **85:**757–761.
37. **Rutherford, T., and A. W. Nienhuis.** 1987. Human globin gene promoter sequences are sufficient for specific expression of a hybrid gene transfected into tissue culture cells. Mol. Cell. Biol. **7:**398–402.
38. **Treisman, R., M. R. Green, and T. Maniatis.** 1983. *cis* and *trans* activation of globin gene transcription in transient assays. Proc.

Natl. Acad. Sci. USA **80**:7428–7432.

39. **Tuan, D., W. Solomon, Q. Li, and I. M. London.** 1985. The "β-like globin" gene domain in human erythroid cells. Proc. Natl. Acad. Sci. USA **82**:6384–6388.
40. **Tuan, D., A. Abeliovich, M. Lee-Oldham, and D. Lee.** 1987. Identification of regulatory elements of human β-like globin genes, p. 211–220. *In* G. Stamatoyannopoulos and A. W. Nienhuis (ed.), Developmental control of globin gene expression. Alan R. Liss, New York.
41. **Venkatesh, L. K., and G. Chinnadurai.** 1987. Activation of the adenovirus 2 protein IX promoter by DNA replication in a transient expression assay. Nucleic Acids Res. **15**:2235–2250.
42. **Whitelaw, E., and N. J. Proudfoot.** 1983. Transcriptional activity of the human pseudogene ψ α globin compared with α globin, its functional gene counterpart. Nucleic Acids Res. **11**:7717–7733.
43. **Wong, T., and E. Neumann.** 1982. Electric field mediated gene transfer. Biochem. Biophys. Res. Commun. **107**:584–587.
44. **Wright, S., E. deBoer, F. G. Grosveld, and R. A. Flavell.** 1983. Regulated expression of the human β-globin gene family in murine erythroleukaemia cells. Nature (London) **305**:333–336.
45. **Yagi, M., R. Gelinas, J. T. Elder, M. Peretz, T. Papayannopoulou, G. Stamatoyannopoulos, and M. Groudine.** 1986. Chromatin structure and developmental expression of the human α-globin cluster. Mol. Cell. Biol. **6**:1108–1116.