MOLECULAR AND CELLULAR BIOLOGY, Nov. 1990, p. 5958–5966
0270-7306/90/115958-09$02.00/0

Vol. 10, No. 11

Promoter Elements and Erythroid Cell Nuclear Factors That Regulate α-Globin Gene Transcription In Vitro

CHUL G. KIM, STEVEN L. SWENDEMAN, KERRY M. BARNHART, AND MICHAEL SHEFFERY*

DeWitt Wallace Research Laboratory and Graduate Program in Molecular Biology, Memorial Sloan-Kettering Cancer Center, New York, New York 10021

Received 3 July 1990/Accepted 14 August 1990

We have previously purified four factors (α-IRP, α-CP1, α-CP2, and NF-E1) that interact with the promoter of the α-globin gene. One of these (NF-E1) is a tissue-restricted factor that has recently been cloned. The binding sites of these factors identify DNA sequence elements that might mediate the tissue-specific and inducible transcription of the α-globin gene. This possibility was tested in a series of in vitro transcription experiments. An examination of 5′ truncated templates and synthetic promoters constituted from individual factor-binding sites apposed to the α-TATAA box showed that the binding elements of three factors (α-CP1, α-IRP, and NF-E1) mediate four- to sixfold activation of transcription in vitro. In contrast, one element (α-CP2) stimulated transcription less than twofold. The 5- to 10-fold stimulation of these latter templates upon addition of a DNA sequence affinity-purified factor suggests that α-CP2 is functionally limiting in nuclear extracts. Additional experiments further tested the effect of supplementing extracts with factors purified from erythroid cell nuclear extracts or, in the case of NF-E1, enriched from a bacterial cDNA expression system. Each factor tested stimulated transcription in vitro in a binding-site-dependent manner. Our results provide a comprehensive functional view of the murine α-globin promoter and suggest possible mechanisms for activation of α-globin gene transcription during induced differentiation of murine erythroleukemia cells.

Our understanding of *cis*-acting elements and *trans*-acting factors that interact coordinately to regulate tissue-specific gene transcription has expanded greatly in the past several years. The description of single- and multiple-factor-binding sites has partially explained differences between weak and strong promoters, enhancer function, and the nature of tissue-specific and inducible gene transcription (for a review, see reference 20). Transcriptional induction of the tissue-specific α-globin gene during erythroid differentiation in murine erythroleukemia (MEL) cells (27) provides a well-defined model for the investigation of how multiple nuclear factors interact to mediate transcriptional activation. Within 48 h after culture in the presence of an inducing agent, α-globin gene transcription increases 10- to 20-fold in MEL cells (27) and α-globin mRNA content increases 100-fold (32). This potent activation of α-globin gene expression is thought to involve, in part, genetic elements within the promoter (32).

We have previously described the purification of four factors that interact with the α-globin promoter (3, 15). Figure 1 illustrates how the factor-binding sites are arrayed. Two sequences that form a pair of inverted repeats (IRs) upstream of the TATAA box bind a polypeptide, α-IRP, with an M_r of 85,000. Further upstream, two factors, α-CP1 and α-CP2, have binding sites that partially overlap in the α-globin CCAAT box (2). We have previously shown that α-CP1 is composed of at least seven polypeptides with M_rs between 27,000 and 38,000 that are organized into a heterotypic complex that can dissociate into at least two components, both of which are required for DNA binding (15). In contrast, α-CP2 comprises a polypeptide doublet with M_rs of 64,000 and 66,000 (15). Purified α-CP2 protects an extended region that spans from −85 (where it overlaps with the binding site of α-CP1) to −160. Upstream from the α-CP2-binding site is a DNA sequence element (5′-GATAAGGA-3′) found near many erythroid-cell-specific promoters (6, 9, 13, 22, 23). This sequence binds a factor, prominent in erythroid cells, that has recently been cloned (13, 30). We have termed this factor EF-1, but it is also known as GF-1, Eryf-1, and NF-E1. In the remainder of this report, we shall refer to this factor as NF-E1.

The factors described above presumably play a role in mediating both tissue-specific transcription and inducible transcription of the α-globin gene. To test this possibility, we initiated a series of in vitro transcription experiments that tested the functional role of the purified factors and each factor-binding site.

MATERIALS AND METHODS

Nuclear extracts. MEL and HeLa cells were maintained and nuclear extracts were prepared as described previously (8, 10, 26), except that all buffers contained aprotinin (1 μg/ml), benzamidine (0.5 mM), bestatin (5 μg/ml), leupeptin (2 μg/ml), pepstatin A (1 μg/ml), and phenylmethylsulfonyl fluoride (0.5 mM). For some experiments, HeLa cell extracts were kindly provided by Joe Tantravahi and Chi-Gun Lee.

Electrophoretic gels. Electrophoretic gel shift assays were performed as previously described (14, 28), and the probe DNAs used are described below. Electrophoresis in sodium dodecyl sulfate (SDS)-polyacrylamide gels was performed as previously described (16), and separated proteins were visualized by silver staining (ICN Pharmaceuticals Inc.). M_rs were determined by reference to those of commercially obtained proteins.

MS protection. Dimethyl sulfate (DMS) protection assays were performed as described previously (8), with DNA sequence affinity-purified α-CP2 (15).

Oligonucleotides. The following synthetic oligonucleotides were prepared on an Applied Biosystems DNA synthesizer: α-IRP (includes both IRs; specific for α-IRP), 5′-GAT CTG CTC CAA GGG CGT GTC CAC CCT GCC TGG AGG ACA

* Corresponding author.

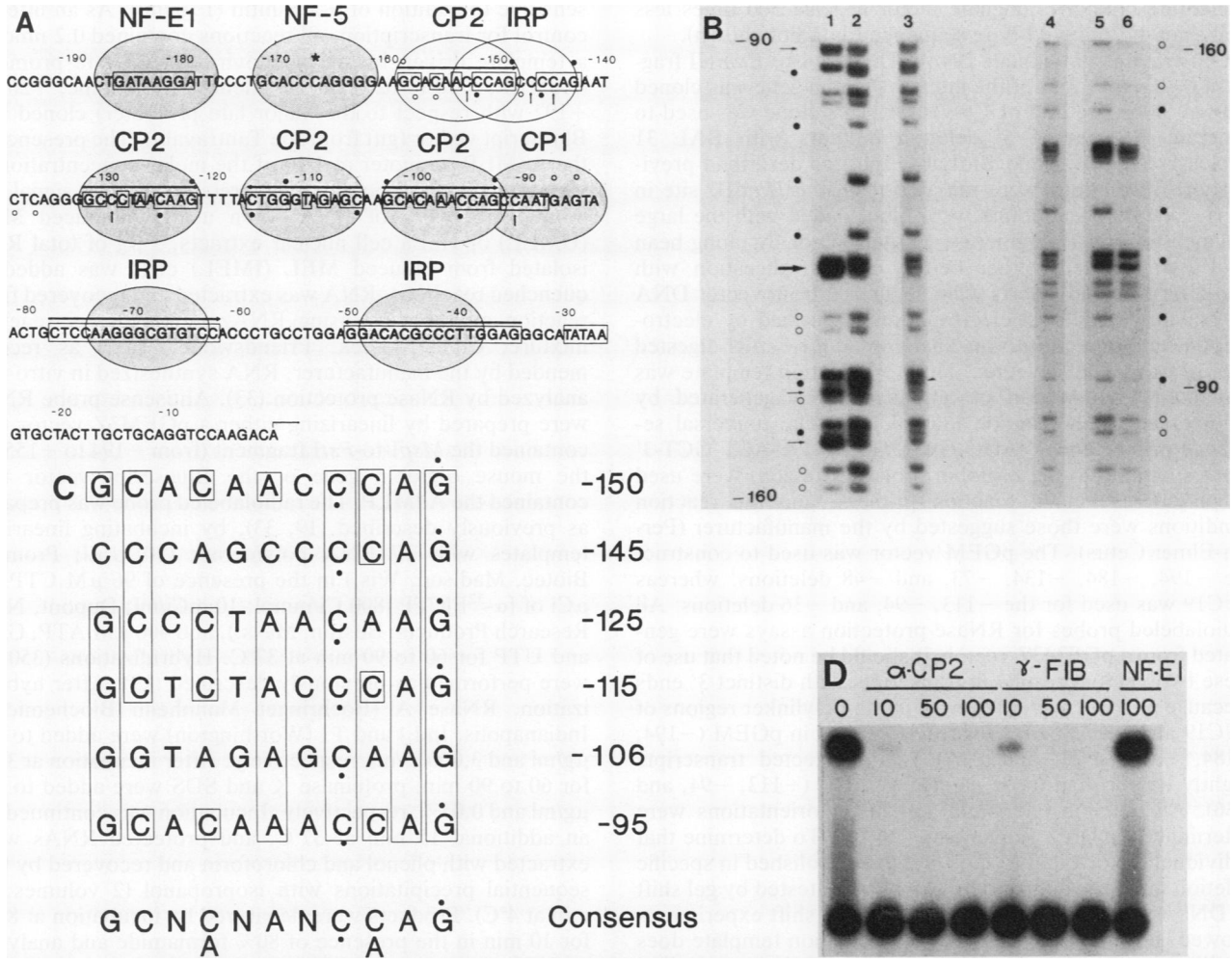

FIG. 1. Factor-binding sites on the α-globin promoter and DMS protection by α-CP2. (A) Schematic representation of factor-binding sites on the α-globin promoter. Shaded ovals represent factors with strong binding sites. Unshaded ovals represent weak binding sites. Various promoter elements are boxed. These include the TATAA element, the IRs that bind α-IRP, the CCAAT box, the consensus binding sequence for α-CP2, and the binding site for NF-E1. The asterisk in the center of the α-NF-5-binding site indicates that mutation of this base (from C to T) abolishes factor binding. Bases protected from methylation by α-CP2 are also illustrated, as described below. (B) DMS protection afforded by DNA sequence affinity-purified α-CP2. Lanes: 1 and 6, α-CP2 added; 2 and 5, no added protein; 3 and 4, Maxam-Gilbert G reaction. These data are displayed on the α-globin promoter in panel A. Protection from methylation is indicated by filled (strong protection) and unfilled (weak protection) circles. The positions of circles above and below the line indicate protection of the sequence as written or its complement, respectively. DMS-hypersensitive sites are indicated by arrows (all hypersensitive sites are on the complementary strand). The numbers to the sides indicate distance from the cap site. (C) Comparison of sequences protected by α-CP2. Strongly protected bases are indicated as described above, and a consensus site (four of six sites) is indicated at the bottom. (D) An oligonucleotide spanning sequences from positions −93 to −125 was labeled and incubated with purified α-CP2 in the presence of the indicated amounts (in nanograms) of the same unlabeled oligonucleotide (designated α-CP2), an oligonucleotide corresponding to the γ-fibrinogen CCAAT box (γ-FIB), or the NF-E1-binding site (NF-E1).

CGC CCT TGG AGG G-3′–5′-AAT TCC CTC CAA GGG CGT GTC CTC GAG GCA GGG TGG ACA CGC CCT TGG AGC A-3′; α-CP1 (specific for α-CP1), 5′-AGC TCA AAT TAA CCA ATC AGC GCA CTC TCA CAG G-3′–5′AGC TCC TGT GAG AGT GCG CTG ATT GGT TAA TTT G-3′; mutα-CP1 (CCAAT→CACAT), 5′-CAC AAA CCA GCA CAT GAG TAA CTG CTC CAA GGG C-3′–05′GCC CTT GGA GCA GTT ACT CAT GTG CTG GTT TGT G-3′; α-CP2 (specific for α-CP2), 5′-GAT CCC AAG TTT TAC TCG GTA GAG CAA GCA CAA ACC AGG-3′–5′-GAT CCC TGG TTT GTG CTT GCT CTA CCG AGT AAA ACT TGG-3′; α-NF-5, 5′-TTA AGG ATT CCC TGC ACC CAG GGG AAG CAC AA-3′–5′-TTA ATT GTG CTT CCC CTG GGT GCA GGG AAT CC-3′; mutα-NF-5 (CACCC→CA CCT), 5′-TTA AGG ATT CCC TGC ACC TAG GGG AAG CAC AA-3′–5′-TTA ATT GTG CTT CCC CTA GGT GCA GGG AAT CC-3′, NF-E1, 5′-GAT CCA ACT GAT AAG GAT TC-3′–5′TTA AGA ATC CTT ATC AGT TG-3′, mutNF-E1, 5′-GAT CCA ACT TCT AAG GAT TC-3′–5′-TTA AGA ATC CTT AGA AGT TG-3′ γ-fibrinogen, 5′-GAT CTG ACC AGT TCC AGC CAC TCT T-3′–5′-GAT CAA GAG TGG CTG GAA CTG GTC A-3′ (5). All α-globin sequences correspond to the DNase I footprint of the purified factor on the α-globin promoter, with the exception of α-NF-5, whose footprint has not been mapped, and α-CP1, which is a sequence derived from the α-globin promoter that maximizes α-CP1 binding and abolishes interactions with α-CP2 (5, 15a). However, the same results were obtained with an oligonucleotide that corresponds exactly to the α-globin promoter (data not shown). Each mutant oligo-

nucleotide bound its cognate factor at least 500 times less well than did the wild-type sequence (data not shown).

Construction of template DNAs. The *Msp*I-to-*Bam*HI fragment (−194 to +155) of the mouse α-globin gene was cloned into an *Sma*I site of a pUC9 vector. This clone was used to generate a series of 5′ deletion mutants with BAL 31 nuclease (New England BioLabs, Inc.) as described previously (8). Deletions were initiated from the *Hin*dIII site in pUC. Digestion endpoints were blunt ended with the large fragment of DNA polymerase I (Klenow) or by mung bean nuclease treatment either before or after digestion with *Eco*RI. Truncated inserts were separated from vector DNA by polyacrylamide gel electrophoresis, purified by electroelution, and ligated into an *Sma*I- or *Sma*I-*Eco*RI-digested pUC19 or pGEM3Z vector. The −94 deletion template was generated by insertion of amplified DNA generated by polymerase chain reaction into pUC19. The universal sequence primer and 5′-AGC CAA TGA GTA ACT GCT-3′ (corresponding to the α-globin promoter region) were used as polymerase chain reaction primers, and the reaction conditions were those suggested by the manufacturer (Perkin-Elmer/Cetus). The pGEM vector was used to construct the −194, −184, −134, −73, and −48 deletions, whereas pUC19 was used for the −113, −94, and −36 deletions. All radiolabeled probes for RNase protection assays were generated from a pGEM3Z vector. It should be noted that use of these two vectors results in transcripts with distinct 3′ ends (because of sequence differences in the polylinker regions of pUC19 and pGEM3Z). Constructs carried in pGEM (−194, −184, −134, −73, and −48) yield protected transcripts slightly longer than those carried by pUC (−113, −94, and −36). All deletion endpoints and insert orientations were determined by DNA sequencing (18, 25). To determine that individual factor-binding sites had been abolished in specific deletion mutants, appropriate clones were tested by gel shift or DNase I footprinting experiments. Gel shift experiments showed, for example, that the −184 deletion template does not bind purified NF-E1. Footprinting experiments showed that the −94 deletion has a single α-CP1-binding site and two α-IRP-binding sites, while the −48 and −73 deletions have one and two α-IRP-binding sites, respectively (data not shown).

Multimerized factor-binding sites were constructed by using appropriate complementary oligonucleotides which were annealed, phosphorylated, and self-ligated to generate multimers as previously described (15). The multimers were blunt ended with Klenow and ligated to a dephosphorylated and blunt-ended *Xba*I site of the −36 deletion mutant. The orientations and copy numbers of the binding-site inserts were determined by DNA sequencing (25).

Transcription in vitro. In vitro transcription reactions were performed with 2 nM supercoiled template DNA and 10 to 15 μl of nuclear extracts (at 10 to 15 mg/ml) (4) in a 50-μl reaction mixture by using a modification of the reaction mixture previously described (29). Template response to inclusion of various upstream elements is best observed when supercoiled templates are used. Standard incubation mixtures contained 10 mM Tris (pH 7.9), 10% (vol/vol) glycerol, 48 mM KCl, 0.3 mM dithiothreitol, 0.5 mM EDTA, 1 mM $MgCl_2$, 40 mM $(NH_4)_2SO_4$, 10 mM creatine phosphate, 0.4 U of phosphocreatine kinase, 200 μM ATP, and CTP, GTP, and UTP, each at 120 μM. Transcription reactions were incubated for 60 min at 30°C and terminated by addition of 1 μg of RNase-free DNase I (DPRF grade; Worthington Diagnostics), followed by incubation for 15 min at 37°C. All reactions were dependent upon added template and were sensitive to addition of α-amanitin (1 μg/ml). As an internal control for transcription, all reactions contained 0.2 nmol of a template driven by the adenovirus major late promoter (AdMLP; 5778 to 6231 on the adenovirus genome; −261 to +192 with respect to the major late promoter) cloned in a Bluescript vector (gift from Joe Tantravahi). The presence of the AdMLP promoter at 1/10 of the molar concentration of globin templates accounts for its relatively weak signal. To monitor recovery of RNA when using uninduced MEL (UMEL) or HeLa cell nuclear extracts, 3 μg of total RNA isolated from induced MEL (IMEL) cells was added to quenched reactions. RNA was extracted and recovered from reaction mixtures by using RNAzol (0.8 ml per reaction mixture; Cinna/Biotecx, Friendswood, Tex.) as recommended by the manufacturer. RNA synthesized in vitro was analyzed by RNase protection (33). Antisense probe RNAs were prepared by linearizing either a pGEM3Z vector that contained the *Msp*I-to-*Pst*I fragment (from −194 to +155) of the mouse α-globin gene or the Bluescript vector that contained the AdMLP. The radiolabeled probe was prepared as previously described (19, 33), by incubating linearized templates with T7 RNA polymerase (10 U/μl; Promega Biotec, Madison, Wis.) in the presence of 96 μM CTP, 50 μCi of [α-^{32}P]CTP (800 Ci/mmol; 10 mCi/ml; Dupont, NEN Research Products, Boston, Mass.), and 400 μM ATP, GTP, and UTP for 60 to 90 min at 37°C. Hybridizations (350 μl) were performed as previously described (19). After hybridization, RNase A (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) and T1 (Worthington) were added to 114 μg/ml and 3,000 U/ml, respectively. After incubation at 37°C for 60 to 90 min, proteinase K and SDS were added to 171 μg/ml and 0.05%, respectively. Incubation was continued for an additional 15 min at 37°C, and protected RNAs were extracted with phenol and chloroform and recovered by two sequential precipitations with isopropanol (2 volumes; 20 min at 4°C). Products were denatured by incubation at 85°C for 10 min in the presence of 80% formamide and analyzed on 7% polyacrylamide–7 M urea gels. After electrophoresis, gels were soaked in water for 5 min, dried, exposed to X-ray film, and visualized by autoradiography. Results were quantified by densitometric scanning of the autoradiograms with an LKB laser densitometer or by cutting out individual bands and counting them by liquid scintillation. AdMLP transcripts were used to correct for any internal variation in the amount of transcription in vitro. It should be noted that globin probes produce two protected fragments (of approximately 165 and 128 bases). The smaller fragment corresponds to endogenous, spliced α-globin mRNA present in IMEL cell nuclear extracts (and added exogenously to HeLa and UMEL cell extracts). The size of the larger protected fragment corresponds to the unspliced in vitro transcript initiated at the α-globin cap site and extended into adjacent plasmid sequences. All in vitro transcription experiments were repeated at least three times with at least two independent nuclear extracts. Factor-depleted extracts were prepared as previously described (1, 31), by incubating nuclear extracts with double-stranded oligonucleotides corresponding to the appropriate factor-binding site for 15 min at room temperature before initiating transcription reactions. Optimum conditions (50 ng of oligonucleotide for 15 μl of extract) were determined empirically by titration.

Bacterial expression and enrichment of NF-E1. A cDNA clone for NF-E1 was isolated from a pCDM8 library (K. M. Barnhart and M. Sheffery, unpublished data). The clone is identical to that described by Tsai et al. (30), except that the 5′ end has an additional 41 bases: 5′-CAG AGC CAA GGC

CAG TGA GGA CTC CCT TGG GAT CAC CCT GA-3'. The cDNA insert was released from the vector and cut at an internal *Nco*I site that overlaps the ATG start codon (30). This fragment was blunt ended by incubation with Klenow and inserted into a T7 polymerase-driven expression vector, pET3c (24), blunt ended at its *Nde*I site. Host strain BL21DE3, which carries the T7 RNA polymerase gene under *lacUV5* control, was transformed, and clones with inserts in the appropriate orientation were identified. Bacterial cultures were grown and induced as previously described (17). To avoid bacterial RNases that inhibit in vitro transcription experiments, NF-E1 was prepared from bacteria by using the denaturation-and-renaturation protocol described by Citovsky et al. (7).

RESULTS

Factor-binding sites on the murine α-globin promoter. Our model of the α-globin promoter (Fig. 1A) incorporates several new details which must be briefly noted. Previous results demonstrated that affinity-purified α-CP2 protects an extended region of the α-globin promoter from digestion by DNase I (2). We determined the nucleotides involved in α-CP2 binding by performing DMS protection experiments. The results (Fig. 1B) indicate that the affinity-purified factor protects a tetranucleotide motif (CCAG or CAAG) repeated six times between −92 and −160 on the α-globin promoter. Comparison of flanking sequences suggests a consensus binding site of GCNCNANCCAG (Fig. 1C; protected sites are also indicated in Fig. 1A). Protection of purines in the CCAG quartet by α-CP2 is reminiscent of the interaction of a HeLa cell CCAAT factor, also termed CP2, with its binding sites (5). Indeed, a strong binding site for HeLa CP2, the γ-fibrinogen CCAAT box, competes effectively (within a factor of 2) for the binding of affinity-purified α-CP2 to α-globin sequences (Fig. 1D). By these criteria, affinity-purified factor α-CP2 appears to be related or identical to HeLa cell CCAAT factor CP2.

The results described above show that α-CP2 binds sequences from −92 to −160 under the saturating conditions used for DMS protection experiments. However, an oligonucleotide covering sequences from −134 to −160 interacted with purified α-CP2 only 1/5 to 1/10 as well as sequences near the CCAAT box (−100; data not shown). Indeed, when used to probe nuclear extracts, this upstream oligonucleotide interacted preferentially, although weakly, with α-IRP (about one-fifth as well as sequences near −70; data not shown). Thus, sequences from −134 to −160 appeared to interact with both α-CP2 and α-IRP, although both factors bound weakly (Fig. 1A).

One additional feature of the α-globin promoter is also included in Fig. 1A. An oligonucleotide that covers sequences from −155 to −182 interacts with a factor, which we term α-NF-5, that is present in both MEL and HeLa cell nuclear extracts. Binding is not competed for by any previously identified α-promoter-binding site nor by oligonucleotides containing AP-2 (21) or CAC_3 factor-binding sites (these latter sites share some sequence similarities with the α-NF-5-binding region; data not shown). Although α-NF-5 has not been well characterized, the five factors illustrated in Fig. 1A (α-IRP, α-CP1, α-CP2, α-NF-5, and NF-E1) appear to account for most factors that interact strongly with the α-globin promoter.

Transcription of truncated α-globin templates in vitro. To assess the functions of the binding sites and nuclear factors described above, we analyzed transcription of several α-globin promoter templates in vitro. Each template was examined for its activity in nuclear extracts prepared from UMEL, IMEL, and HeLa cells. As an internal control, a template driven by the AdMLP was included in all of the reaction mixtures (at a molar template [AdMLP/globin] ratio of 1:10). Transcripts were assayed by RNase protection (see Materials and Methods). Note that IMEL cell extracts contain significant amounts of endogenous α-globin mRNA which facilitate quantification of RNA recovery. In contrast, UMEL and HeLa cell extracts lack substantial amounts of endogenous α-globin mRNA. To monitor RNA recovery in these instances, mRNA prepared from IMEL cell extracts was usually added to quenched in vitro transcription reactions performed with UMEL or HeLa cell nuclear extracts. Thus, in addition to the use of AdMLP as an internal control for in vitro transcription, most reaction mixtures also contained added or endogenous α-globin mRNA as an internal control for RNA recovery.

First, we tested a series of α-globin promoter deletion templates for the ability to be transcribed in vitro by IMEL, UMEL, or HeLa cell nuclear extracts. The results (Fig. 2) showed that transcription from the α-globin promoter was strongly dependent on upstream sequences. For example, promoters that comprise TATAA alone (−36) were expressed at levels from 4 to 15% of that of the wild-type template (−194; no products could be detected when templates without a TATAA box were assayed). Progressive addition of upstream sequences progressively stimulated in vitro transcription. For example, while templates with one α-IRP-binding site were stimulated less than twofold, templates with two α-IRP-binding sites were stimulated fourfold compared with TATAA alone. Addition of the α-CP1-binding site (the α-globin CCAAT box) resulted in further stimulation of transcription in all three extracts. Templates that extend progressively further upstream were stimulated further. The quantitative differences were relatively small, however, and it is difficult to assess accurately the specific contributions made by factors bound to sites more distal than −94 (the CCAAT box). In addition, there were significant standard errors (15 to 20%) in results obtained from different experiments. Because of these limitations in interpretation, we constructed a series of synthetic promoter templates to dissect further the functions of individual factor-binding sites in vitro.

Multimerized promoter elements. To examine each promoter element in isolation, we constructed a series of templates composed of multimerized factor-binding sites appended (at −36) to the α-globin TATAA box. Figure 3D illustrates the structure of each template and shows the results obtained when individual templates were transcribed in the three nuclear extracts. A dimer of the IR element ligated to the TATAA box (IR/TATAA) resulted in three- to fourfold stimulation of activity compared with TATAA alone, about the same stimulation observed with deletion templates which contain two IRs. The results in Fig. 3 show additionally that the two IR elements function in the absence of linker sequences and in an orientation distinct from their natural context.

Multimers of the α-CP1-binding site (α-CP1/TATAA) also directly stimulated TATAA-driven transcription, approximately fivefold in this case (Fig. 3). In contrast, multimerization of the α-CP2-binding site (α-CP2/TATAA) had a relatively small direct effect on α-globin transcription in vitro, and templates containing multimerized α-CP2-binding sites were activated less than twofold in all extracts. The poor responses of these templates might be explained in

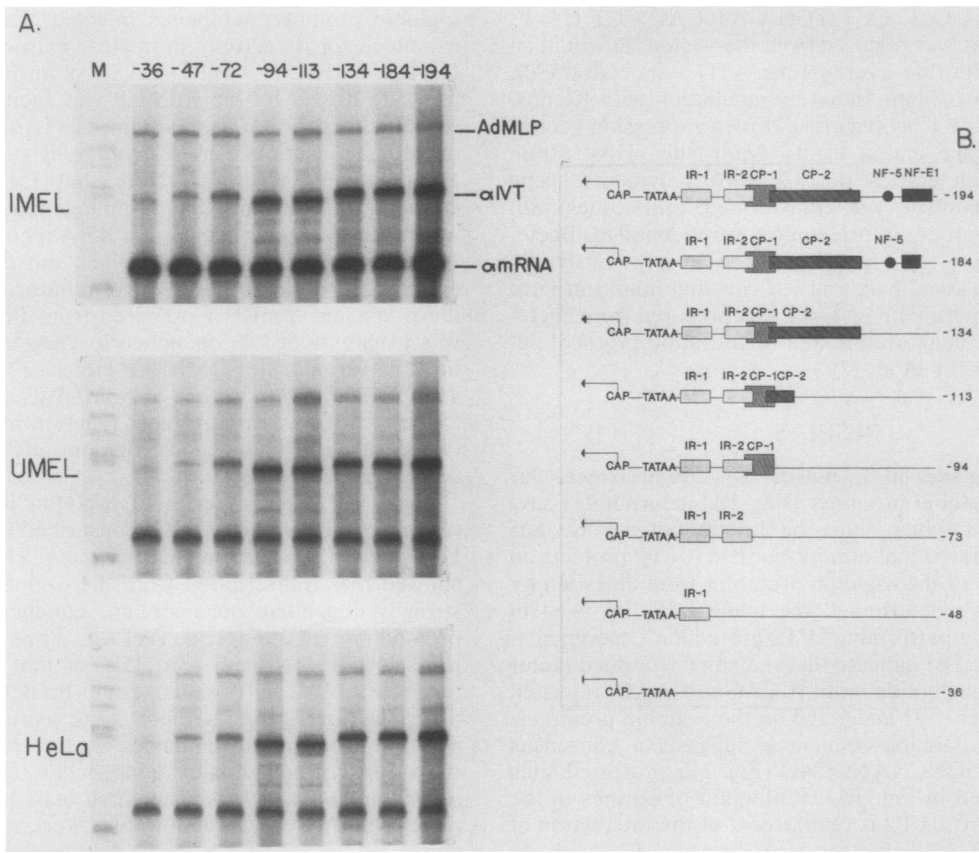

FIG. 2. Transcription of α-globin deletion templates in IMEL, UMEL, and HeLa cell nuclear extracts. (A) Templates that extend to the nucleotide shown at the top of the panel (see Fig. 1 for map details) were transcribed in extracts prepared from UMEL, IMEL, or HeLa cells, as indicated. The mobilities of probe RNAs protected by transcripts initiated in vitro from the AdMLP or α-globin (αIVT) promoter or by α-globin mRNA are indicated to the right of the IMEL reactions. Lane M contained molecular size markers. (B) Schematic representations of factor-binding sites deleted in the constructs used.

several ways. For example, α-CP2 might be limiting under these experimental conditions or α-CP2 might have to interact with neighboring factors to influence α-globin gene transcription. In considering these two (nonexclusive) possibilities, we favor the former explanation. As shown below, addition of affinity-purified α-CP2 to reaction mixtures containing α-CP2/TATAA templates significantly stimulated their transcription in vitro.

Multimerized NF-E1-binding sites also constitute strong erythroid cell-specific promoter elements. Addition of NF-E1 sites directly stimulated the TATAA element five- to sixfold in vitro (Fig. 3). As expected, the NF-E1/TATAA template behaved like TATAA alone when assayed in nuclear extracts prepared from HeLa cells, which are devoid of NF-E1 activity (Fig. 3A and C, compare lanes marked NF-E1).

Addition of purified factors. To further examine the roles that individual factors play in regulating α-globin gene transcription in vitro, we supplemented crude or factor-depleted extracts with various purified or enriched proteins. To simplify interpretation, we used the multimerized factor-binding templates described above.

Addition of affinity-purified α-CP1 (Fig. 4B) to in vitro transcription reaction mixtures containing the α-CP1/TATAA promoter did not significantly influence its activity (Fig. 4A, compare lanes 2 and 3). Since addition of α-CP1-binding sites strongly stimulates TATAA-driven transcription, we suspected that the poor response to the added factor might be due to the presence of saturating amounts of α-CP1 in nuclear extracts. To test this possibility, we depleted extracts of free α-CP1 by incubating them in the presence of an oligonucleotide specific for this factor. After depletion, there was a threefold drop in the transcriptional activity of the α-CP1/TATAA template (Fig. 4A, lane 4). As expected, activity was restored by addition of excess affinity-purified α-CP1 to the depleted extracts (Fig. 4A, lane 5). To control for nonspecific effects, the same experiment was performed with a template composed of a multimerized mutant α-CP1 oligonucleotide (mutα-CP1/TATAA). The mutant oligonucleotide, which binds α-CP1 1,000-fold less well than the wild-type oligonucleotide (15a), is present in the same copy number and orientation as the wild-type oligonucleotide in the α-CP1/TATAA template (data not shown). The mutant template was expressed about one-fourth as well as the wild-type template in crude extracts (Fig. 4A), a level similar to that of templates that include TATAA alone (Fig. 2 and 3). The activity of the mutα-CP1/TATAA template was not affected by depletion of extracts of α-CP1, nor was it increased by addition of purified α-CP1 to crude or factor-depleted extracts. These results suggest that α-CP1 acts directly through its binding site on TATAA-bound factors to stimulate α-globin gene transcription. In addition, our results suggest that α-CP1 is present at nearly saturating levels in IMEL cell extracts.

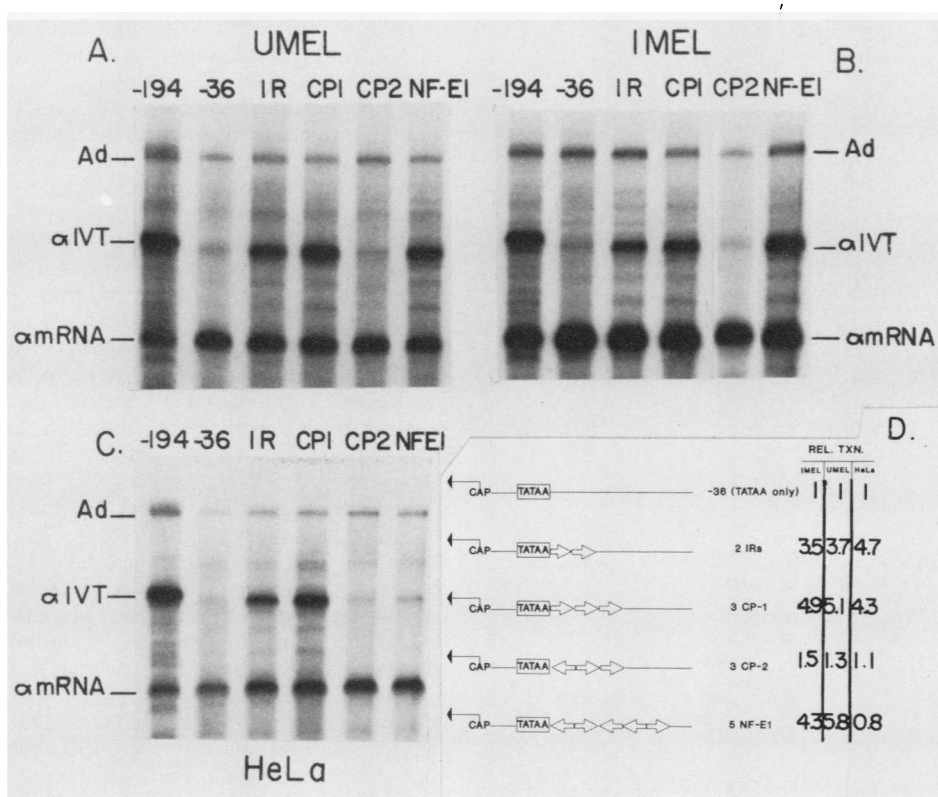

FIG. 3. Action of individual factor-binding sites as stimulatory promoter elements. Templates that extend to position −194 or −36 (the α-globin TATAA box) or are constituted from the indicated multimerized factor-binding site ligated to the −36 template were transcribed in vitro in UMEL (A), IMEL (B), or HeLa (C) cell nuclear extracts. The positions of transcripts initiated in vitro from the AdMLP (Ad) or α-globin (αIVT) templates and that of α-globin mRNA (αmRNA) are indicated at the side of each panel. The structure of each multimerized template is illustrated in panel D. REL. TXN., Relative transcription. The fold activation of each multimerized template relative to TATAA alone is indicated to the right of panel D.

Although α-CP1 appears to be present near saturation in IMEL cell extracts, the relatively poor activity of the multimerized α-CP2/TATAA templates described above suggested that α-CP2 was present in limiting amounts in nuclear extracts. Under these circumstances, addition of affinity-purified α-CP2 to crude extracts should directly stimulate transcription of α-CP2/TATAA templates. Indeed, affinity-purified α-CP2 (Fig. 5D) stimulated the α-CP2/

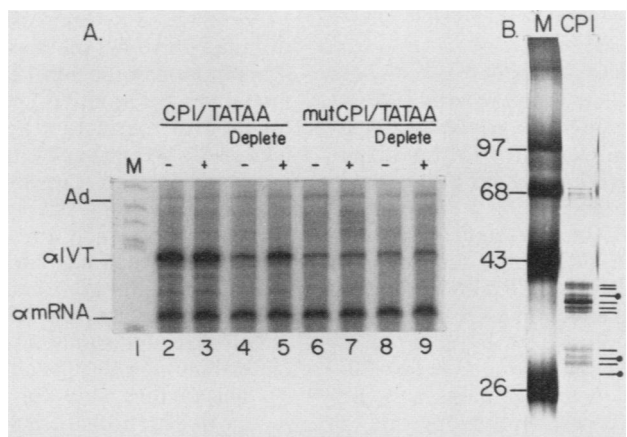

FIG. 4. In vitro transcription stimulated by purified α-CP1. (A) A template constituted from an α-CP1-binding site multimerized and ligated to the α-globin TATAA box (Fig. 3) or a mutant α-CP1-binding site (see Materials and Methods) was incubated with (+) or without (−) DNA sequence affinity-purified α-CP1 (see panel B). Extracts were depleted of α-CP1-binding activity as indicated (see Materials and Methods). Lane M contained DNA size markers. (B) DNA sequence affinity-purified α-CP1 used in the transcription assays. We have previously shown that affinity-purified α-CP1 is composed of at least seven polypeptides with M_rs of 27,000 to 38,000. The positions of these polypeptides are indicated by lines to the right. Bands indicated by lines ending in closed circles are polypeptides that were sometimes seen in preparations of α-CP1 and might represent proteolytic fragments. Lane M contained size marker proteins whose M_rs (10^3) are indicated to the left.

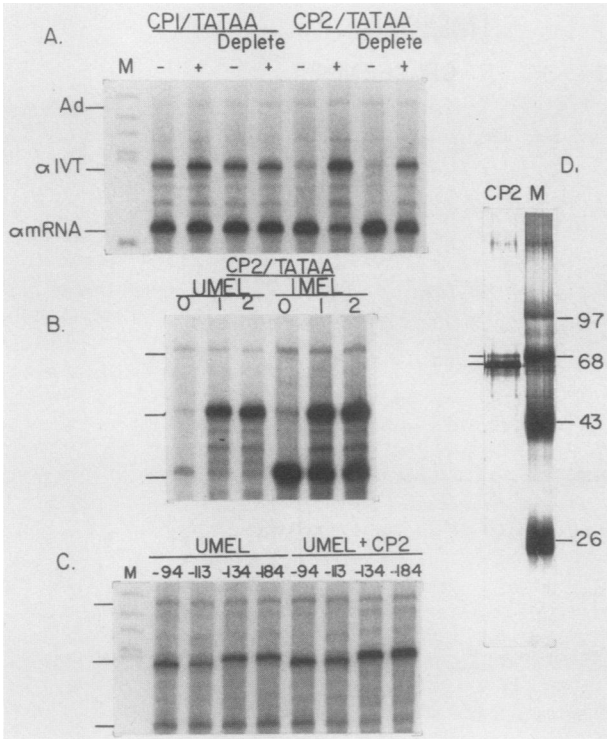

FIG. 5. Transcriptional activation by purified α-CP2. (A) Templates constituted from α-CP1 (left)- or αCP2 (right)-binding sites ligated to the α-globin TATAA box (Fig. 3) were incubated without (−) or with (+) DNA sequence affinity-purified α-CP2. Extracts were depleted of α-CP2-binding activity as indicated. (B) Comparison of α-CP2/TATAA templates in UMEL and IMEL extracts without (left) or with (right) the indicated amount (in microliters) of purified α-CP2. The template used was a single copy of an oligonucleotide that spans positions −93 to −125 ligated to the TATAA box in an orientation opposite from that in the natural promoter. (C) Templates that extend to the indicted nucleotides were transcribed in nuclear extracts prepared from UMEL cells without (left) or with (right) added α-CP2 (2 μl). (D) DNA sequence affinity-purified α-CP2 used in the in vitro transcription assays. Previous results have shown that α-CP2 is composed of a polypeptide doublet with M_rs of 64,000 and 66,000. These polypeptides are indicated by horizontal lines to the left. Lane M contained marker proteins whose sizes are indicated as in Fig. 4.

TATAA promoter 3- to 10-fold when added to crude MEL cell extracts (Fig. 5A and B). The effect was specific, since α-CP1/TATAA templates were unresponsive to the added factor (Fig. 5A). Depletion of α-CP2 had little effect on α-CP1/TATAA or α-CP2/TATAA template expression consistent with the notion that α-CP2 was present in limiting amounts in our nuclear extracts. Addition of excess purified α-CP2 to depleting extracts stimulated α-CP2/TATAA templates at least threefold but had no effect on α-CP1/TATAA.

Because of the strong stimulation caused by addition of purified α-CP2 to a template containing multiple α-CP2-binding sites, we tested a template that contains a single insert of the factor-binding site in an orientation opposite to that found in the natural promoter. This individual element also responded strongly (10-fold) to addition of the purified factor (Fig. 5B). We also tested the effect of adding purified α-CP2 to α-globin templates, described above, that progressively include sequences from the CCAAT box up to but not encompassing the NF-E1-binding site. The transcriptional activity of each template was assayed in UMEL cell nuclear extracts and in extracts supplemented with affinity-purified α-CP2. Addition of affinity-purified α-CP2 to UMEL cell extracts resulted in averages of 1.7- and 2.5-fold stimulation of templates that extend to −113 and −134, respectively (Fig. 5C). In contrast, no further stimulation was manifested by templates that extend to −184, even when twice as much of the purified factor was added (data not shown). Thus, the functional sequences for α-CP2 in the intact α-globin promoter appear to extend to −134 but not to −184.

Lastly, we investigated the role that NF-E1 might play in directly stimulating α-globin gene transcription in vitro. For these experiments, we took advantage of the fact that HeLa cell extracts are naturally devoid of NF-E1 binding activity. In addition, we used as our source of NF-E1 a bacterially synthesized protein. A cDNA clone for murine NF-E1 has recently been described (30), and we have also isolated a cDNA clone for this factor. The cDNA insert, cloned in eucaryotic expression vector pCDM8, is identical to that obtained by Tsai et al. (30), except that 41 additional bases are included in the 5′ nontranslated region (see Materials and Methods). The cDNA that encodes NF-E1 was transferred into an isopropyl-β-D-thiogalactopyranoside-inducible procaryotic expression vector (24). Figure 6A and B shows SDS and electrophoretic shift gels of uninduced and induced bacterial cultures. Induction resulted in the appearance of a major new band with an M_r of 43,000 in SDS-polyacrylamide gels, consistent with the known mass of the cloned factor but slightly smaller than its reported M_r (30). Bacterial lysates contained a DNA-binding activity specific for the NF-E1-binding site only when cultures were grown in the presence of isopropyl-β-D-thiogalactopyranoside (Fig. 6B, compare lanes 2 and 3). Addition of bacterial NF-E1 to HeLa cell nuclear extracts resulted in fivefold stimulation of templates containing NF-E1-binding sites (Fig. 6C). Templates without NF-E1-binding sites, in contrast, were unaffected by addition of the bacterially produced factor. In addition, templates assayed in HeLa cell extracts supplemented with uninduced bacterial lysates were unresponsive (Fig. 6C). We conclude that NF-E1, like α-CP1 and α-CP2, acts directly to stimulate transcription of the α-globin promoter in vitro.

DISCUSSION

We initiated a series of in vitro experiments designed to examine how multiple purified erythroid cell factors mediate transcription of the murine α-globin gene. Having refined our model of the α-globin promoter (Fig. 1), we tested the function of the factors and factor-binding sites previously identified. We began by testing the behavior of a series of truncated promoters in vitro. Progressive inclusion of factor-binding sites progressively increased transcription in vitro, suggesting that each upstream factor-binding element and, presumably, each cognate factor act to stimulate transcription of the α-globin promoter in vitro (Fig. 2).

To examine the role of each factor-binding site in more detail, we tested synthetic promoters composed of individual factor-binding sites ligated to the α-globin TATAA box (Fig. 3). The binding sites for α-IRP and the α-globin CCAAT factor, α-CP1, both mediated three- to fivefold stimulation of α-globin gene transcription. Consistent with this observation, deletions that remove these binding sites significantly reduced promoter activity in all of the extracts tested (Fig. 2). In addition, α-CP1 appears to be an abundant factor that was present in nearly saturating amounts in our nuclear extracts. Indeed, to show any direct effect of addition of the purified factor, α-CP1 activity had to be functionally de-

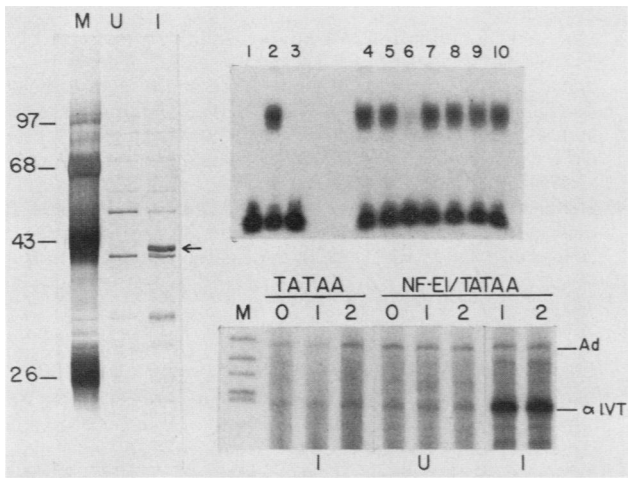

FIG. 6. Direct stimulation of an NF-E1-binding-site-containing promoter by addition of bacterial NF-E1 to HeLa cell extracts. (A) Bacteria containing an isopropyl-β-D-thiogalactopyranoside-inducible NF-E1 expression vector were cultured without (U) or with (I) isopropyl-β-D-thiogalactopyranoside, lysates were prepared as described in Materials and Methods, and proteins were separated on an SDS-polyacrylamide gel and visualized by silver staining. The mobilities of marker proteins (M) are indicated to the left, as in Fig. 4. An induced protein with an M_r of 43,000 is indicated by an arrow to the right. Samples of the material illustrated were used to generate the data shown in panels B and C. (B) Gel shifts of bacterial NF-E1. The probe, composed of sequences from −176 to −189 on the α-globin promoter (and specific for NF-E1) was incubated in the presence of 1 μg of poly(dI-dC) (except lane 4, which was incubated with no competitor), protein, and specific competitors, as indicated. Lanes: 1, probe alone; 2, induced bacterial lysate; 3, uninduced bacterial lysate; 4 to 10, induced bacterial lysate with no competitor (lane 4), 1 μg of poly(dI-dC) (lane 5), 100 ng of NF-E1 oligonucleotide (lane 6), 100 ng of mutNF-E1 (lane 7), 100 ng of α-IRP (lane 8), 100 ng of α-CP1 (lane 9), or 100 ng of α-CP2 (lane 10). The electrophoretic shift generated by the bacterial protein is indistinguishable from that generated by the NF-E1 present in MEL cell nuclear extracts. (C) The indicated amounts (in microliters) of NF-E1 renatured from uninduced (U) or induced (I) bacterial lysates was added to nuclear extracts prepared from HeLa cells. Templates contained either multimerized NF-E1-binding sites ligated to the TATAA box (NF-E1/TATAA) or TATAA alone (TATAA). Lane M contained labeled DNA size markers.

pleted from the transcription assay (Fig. 4). Under these conditions, however, purified α-CP1 was a strong positive regulator of the α-globin promoter.

Of the four factors we have purified, we have not tested the direct effect of adding α-IRP to in vitro transcription reaction mixtures. The observations discussed above, however, strongly suggest that α-IRP also acts to stimulate α-globin gene transcription. In addition, we have previously demonstrated that the nucleotide sequence of the α-IRP-binding site is similar to that of SP1 and that α-IRP binds effectively to SP1 sites on the simian virus 40 21-base-pair repeat (2). These observations suggest a close relationship between α-IRP and HeLa cell transcription factor SP1 (11).

In contrast to the moderately strong in vitro effects of α-IRP- and α-CP1-binding sites, the α-CP2 element appeared, in the absence of an added factor, to be a weak activator of α-globin transcription (Fig. 3 and 5). Addition of purified α-CP2, on the other hand, stimulated promoter activity 5- to 10-fold when the α-CP2-binding site abutted the TATAA element (Fig. 5). These results suggest that under saturating conditions, α-CP2 is comparable in strength to the other transcription factors. In addition, our results suggest that the functional α-CP2-binding domain extends from −94 to −134 (Fig. 5C).

Our results also demonstrate that NF-E1 is a strong tissue-specific transcription factor when assayed in vitro. This conclusion is based on the activation of NF-E1/TATAA promoters in nuclear extracts prepared from MEL cells but not in those prepared from HeLa cells (Fig. 3). In addition, multimerized NF-E1/TATAA templates, which behaved like TATAA alone when assayed in HeLa cell extracts (Fig. 3), were activated by supplementation of the HeLa cell extracts with bacterially synthesized NF-E1 (Fig. 6). In both in vitro assays, the transcriptional activation afforded by NF-E1 appears to be nearly equivalent to that of α-CP1, α-IRP, or α-CP2. In contrast, deletion of NF-E1-binding sites from the intact α-globin promoter results in a relatively small decrement in template activity (20%; Fig. 2). This difference suggests that when NF-E1-binding sites are in the presence of other strong activating elements or situated in a relatively remote location (or both), the direct effects of NF-E1 are significantly diminished in vitro.

Together, our results suggest that two sites (α-IRP and α-CP1) adjacent to the TATAA box bind two constitutively expressed transcription factors abundant in both MEL and HeLa cells. Presumably because of their proximity to the TATAA box, these two factors play a significant role in driving transcription from the intact α-globin promoter in vitro. Indeed, to characterize accurately the role of more distal factor-binding sites, individual elements had to be analyzed in isolation. Upstream from these constitutive factor-binding sites are several elements that interact with a factor (α-CP2) whose binding activity has been shown previously to be induced two- to threefold during erythroid differentiation of MEL cells (2). In addition, the results presented here suggest that α-CP2 is present in functionally limiting amounts compared with α-IRP, α-CP1, and NF-E1 (Fig. 3 and 5). Together, these observations suggest that the induction in α-CP2-binding activity might function in vivo to assist in the induction of α-globin gene transcription during erythroid differentiation. A further test of this hypothesis, however, will require additional in vivo studies.

Further upstream are binding sites for α-NF-5, whose function remains to be tested, and the tissue-restricted transcription factor NF-E1. Our results show clearly that NF-E1 is a potent activator of α-globin gene transcription in vitro, and presumably this tissue-restricted transcription factor plays a key role in activating globin promoters exclusively in erythroid cells.

Our results demonstrate that each of the four factors we have purified are transcriptional activators when assayed in vitro. It appears likely, therefore, that the interplay of these factors in vivo contributes to activation of the α-globin promoter during MEL cell differentiation. Although one of these factors (α-CP2) is slightly induced during differentiation (2), it is not apparent how this small change in factor activity, by itself, accounts for the observed 20-fold increase in α-globin gene transcription. The mechanism that triggers activation of α-globin gene transcription in vivo, therefore, remains to be understood.

ACKNOWLEDGMENTS

This work was supported in part by Public Health Service grants CA-31768 and CA-08748 from the National Cancer Institute and DK-37513 and DK-41079 from the National Institutes of Health and by the Businessmen's Assurance Fund. S.L.S. is a Fellow of the American Cancer Society.

The technical assistance of Renee Ford and Kendra Dean is gratefully acknowledged.

LITERATURE CITED

1. **Andrisani, O. M., Z. Zhu, D. A. Pot, and J. E. Dixon.** 1989. In vitro transcription directed from the somatostatin promoter is dependent upon a purified 43 kDa DNA binding protein. Proc. Natl. Acad. Sci. USA **86**:2181–2185.
2. **Barnhart, K. M., C. G. Kim, S. S. Banerji, and M. Sheffery.** 1988. Identification and characterization of multiple erythroid cell proteins that interact with the promoter of the murine α-globin gene. Mol. Cell. Biol. **8**:3215–3226.
3. **Barnhart, K. M., C. G. Kim, and M. Sheffery.** 1989. Purification and characterization of an erythroid cell-specific factor that binds the murine α- and β-globin genes. Mol. Cell. Biol. **9**:2606–2614.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein using the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
5. **Chodosh, L. A., A. S. Baldwin, R. W. Carthew, and P. A. Sharp.** 1988. Human CCAAT-binding proteins have heterologous subunits. Cell **53**:11–24.
6. **Chretien, S., A. Dubart, D. Beaupain, N. Raich, B. Grandchamp, J. Rosa, M. Goossens, and P.-H. Romeo.** 1988. Alternative transcription and splicing of the human porphobilinogen deaminase gene result either in tissue-specific or in housekeeping expression. Proc. Natl. Acad. Sci. USA **85**:6–10.
7. **Citovsky, V., D. Knorr, G. Schuster, and P. Zambryski.** 1990. The P30 movement protein of tobacco mosaic virus is a single-strand nucleic acid binding protein. Cell **60**:637–647.
8. **Cohen, R. B., M. Sheffery, and C. G. Kim.** 1986. Partial purification of a nuclear protein that binds to the CCAAT box of the mouse α_1-globin gene. Mol. Cell. Biol. **6**:821–832.
9. **deBoer, E., M. Antoniou, V. Mignotte, L. Wall, and F. Grosveld.** 1988. The human β-globin promoter; nuclear protein factors and erythroid-specific induction of transcription. EMBO J. **7**:4203–4212.
10. **Dignam, J. D., P. L. Martin, B. S. Shastry, and R. G. Roeder.** 1983. Eukaryotic gene transcription with purified components. Methods Enzymol. **101**:582–598.
11. **Dynan, W. S., and R. Tjian.** 1983. The promoter specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell **35**:79–87.
12. **Evans, T., and G. Felsenfeld.** 1989. The erythroid-specific transcription factor Eryf1: a new finger protein. Cell **58**:877–885.
13. **Evans, T., M. Reitman, and G. Felsenfeld.** 1988. An erythrocyte-specific DNA binding factor recognizes a regulatory sequence common to all chicken globin genes. Proc. Natl. Acad. Sci. USA **85**:5976–5980.
14. **Garner, M. M., and A. Rezvin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. Nucleic Acids Res. **9**:3047–3060.
15. **Kim, C. G., K. M. Barnhart, and M. Sheffery.** 1988. Purification of multiple erythroid cell proteins that bind the promoter of the α-globin gene. Mol. Cell. Biol. **8**:4270–4281.

15a. **Kim, C. G., and M. Sheffery.** 1990. Physical characterization of the purified, CCAAT transcription factor, α-CP1. J. Biol. Chem. **265**:13362–13369.

16. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
17. **Lin, S.-W., J. J. Dunn, F. W. Studier, and D. W. Stafford.** 1987. Expression of human factor IX and its subfragments in Escherichia coli and generation of antibodies from the subfragments. Biochemistry **26**:5267–5274.
18. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
19. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
20. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371–378.
21. **Mitchell, P. J., C. Wang, and R. Tjian.** 1987. Positive and negative regulation of transcription in vitro: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. Cell **50**:847–861.
22. **Plumb, M., J. Frampton, H. Wainwright, M. Walker, K. Macleod, G. Goodwin, and P. Harrison.** 1989. GATAAG; a cis-control region binding an erythroid-specific nuclear factor with a role in globin and non-globin gene expression. Nucleic Acids Res. **17**:73–92.
23. **Reitman, M., and G. Felsenfeld.** 1988. Mutational analysis of the chicken β-globin enhancer reveals two positive-acting domains. Proc. Natl. Acad. Sci. USA **85**:6267–6271.
24. **Rosenberg, A. H., B. N. Lade, D.-S. Chui, S.-W. Lin, J. J. Dunn, and F. W. Studier.** 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56**:125–135.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
26. **Schreiber, E., P. Matthias, M. M. Muller, and W. Schaffner.** 1989. Rapid detection of octamer binding proteins with 'mini-extracts' prepared from a small number of cells. Nucleic Acids Res. **17**:6419–6420.
27. **Sheffery, M., P. A. Marks, and R. A. Rifkind.** 1984. Gene expression in murine erythroleukemia cells: transcriptional control and chromatin structure of the α_1-globin gene. J. Mol. Biol. **172**:417–436.
28. **Strauss, F., and A. Varshavsky.** 1984. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell **37**:889–901.
29. **Tolunay, H. E., L. Yang, W. M. Kemper, B. Safer, and W. F. Anderson.** 1984. Homologous globin cell-free transcription system with comparison of heterologous factors. Mol. Cell. Biol. **4**:17–22.
30. **Tsai, S.-F., D. I. K. Martin, L. I. Zon, A. D. D'Andrea, G. G. Wong, and S. H. Orkin.** 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. Nature (London) **339**:446–451.
31. **Vaulont, S., N. Puzenat, A. Kahn, and M. Raymondjean.** 1989. Analysis by cell-free transcription of the liver-specific pyruvate kinase promoter. Mol. Cell. Biol. **9**:4409–4415.
32. **Wright, S., E. deBoer, F. G. Grosveld, and R. A. Flavell.** 1983. Regulated expression of the human β-globin gene family in murine erythroleukemia cells. Nature (London) **395**:333–336.
33. **Zinn, K., D. DiMaio, and T. Maniatis.** 1983. Identification of two distinct regulatory regions adjacent to the human β-interferon gene. Cell **34**:865–879.