MOLECULAR AND CELLULAR BIOLOGY, Mar. 1986, p. 821–832
0270-7306/86/030821-12$02.00/0

Vol. 6, No. 3

Partial Purification of a Nuclear Protein That Binds to the CCAAT Box of the Mouse α_1-Globin Gene

ROGER B. COHEN,† MICHAEL SHEFFERY,* AND CHUL GEUN KIM

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

Received 23 July 1985/Accepted 5 December 1985

We enriched a fraction from nuclear extracts of murine erythroleukemia cells which contains a protein able to form stable complexes with the promoter region of the α_1-globin gene. Binding activity, which is present in mouse brain and a variety of cultured mouse and human cell lines, is not erythroid cell specific. Binding studies with α-globin gene promoter deletion mutants as well as DNase I footprinting and dimethyl sulfate protection studies showed that the factor bound specifically to the CCAAT box of the α_1 promoter. Enriched factor preparations exhibited weak binding to the promoter region of the β_{maj}-globin gene. This suggests that this protein could bind differentially to these two promoters in vivo. The enriched factor may be a ubiquitous nuclear protein involved in the differential regulation of the α_1- and β_{maj}-globin genes.

As in other species, the timing of adult α- and β-globin gene expression during mouse development is asynchronous. Adult α genes are expressed during early fetal development (day 7), whereas expression of adult β genes is not detected until the site of hematopoiesis migrates from the blood islands of the yolk sac to the fetal liver some 3 days later (35). Mechanisms that can account for the differential activation of the adult α- and β-globin genes during development have not been well characterized.

Sequence analysis of 5′-flanking regions shows that both α- and β-globin genes share a homologous sequence at approximately −30 (TATAA, or the TATAA box), as they do with several other eucaryotic promoters (4, 13). An intact TATAA box is the minimal requirement for correct transcript initiation from both genes, as determined by in vitro transcription assays (23, 50). In addition, both genes share another widely conserved element, the CCAAT box, which is located at approximately −90 in α and −75 in β (4, 13). Deletions within the CCAAT box dramatically reduce both α- and β-globin promoter usage in transient transfection assays (10, 27, 38).

Despite these similarities, several additional studies reveal that there may be fundamental differences between the α and β promoters. Transient expression assays in nonerythroid cells, for example, show that when sequences upstream of the β-CCAAT box (from −90 to −100) are deleted, β-globin gene expression is reduced 10-fold (10). These studies define an element distal to the CCAAT box that is important for β-globin gene expression. The distal element, which contains the sequence CACCC, is conserved upstream of β-globin genes in many different species, but a homologous element is not present near α-globin promoters (25). Transient expression assays also reveal that β-globin gene expression is readily detected only when the gene is linked to an enhancer (1, 10, 27, 51). Expression of the α-globin gene, on the other hand, does not depend on linkage to a viral enhancer, although enhancers can stimulate α expression (27, 38, 51). Interestingly, the dependence of β-globin gene expression on a *cis*-linked enhancer can be circumvented in *trans* if the gene is transferred into cells which express viral E1a gene products. The effects of E1a gene products, which might also stimulate α-globin gene expression, seem to be mediated through the β-globin TATAA box, without requiring upstream elements (22, 51).

When the human α- and β-globin genes are stably introduced by transfection into murine erythroleukemia (MEL) cells, which can be chemically induced to express erythroid differentiation, the α-globin gene is equally active in uninduced and induced cells. The β-globin gene, however, is expressed only after induction and at a level 100-fold less than that of the α-globin gene (5, 6, 53). The significantly lower expression of the transfected β-globin gene when compared with an α-globin gene again suggests fundamental differences between the two promoters. The inducibility of the β-globin gene, on the other hand, is at least partly due to sequences downstream of the mRNA capping site (6, 54). These results suggest that globin gene promoters and their bodies both play distinctive roles in modulating the tissue-specific expression of these genes.

Additional support for the idea that there are fundamental differences between the α- and β-globin promoters comes from studies of the chromatin structure of the two genes in MEL cells. Results of recent experiments have shown that the extent of nucleosome disruption upstream of the α- and β-globin genes differs (8) and that different patterns of nuclease-sensitive sites are maintained near the two genes, even in the absence of gene transcription (26, 44–46). The distinctive pattern of nuclease-sensitive sites is of particular interest because recent reports indicate that specific nonhistone proteins bind within these sensitive domains (15, 16, 28, 55). These latter results suggest that the distinctive pattern of nuclease-sensitive sites near the α- and β-globin genes in MEL cells may be due to the differential binding of regulatory proteins to the two genes in vivo.

The distinctive enhancer dependence of the β-globin gene in transient transfection assays, the high constitutive expression of α compared with the low and inducible expression of the β-globin gene after stable introduction into MEL cells, and the different patterns of nucleosome structure and nuclease-sensitive sites established near these genes in MEL cells suggest that there are fundamental differences between the two globin promoters. One proposal which accounts for

* Corresponding author.
† Present address: Division of Hematology, Mount Sinai Hospital, New York, NY 10029.

some of the differences in promoter activity and structure described above is that the α promoter has the ability to assemble and elongate transcription complexes using generalized cellular transcription factors. Assembly of active transcription complexes on the β-globin promoter, on the other hand, might require an erythroid-specific factor(s) (6). In support of this latter idea, factors that bind to cloned upstream regions of the chicken β-globin gene in vitro have been identified. Characterization of one factor(s) has shown that it binds to two discrete domains upstream of the chicken β-globin gene and that it is present only in erythroid tissues (15, 16).

MATERIALS AND METHODS

Cell lines and tissue. MEL cells were grown and maintained as described previously (8). For large-scale purification of protein, 10-liter cultures of MEL cells (10^6 cells per ml) were grown in roller bottles or spinner flasks. NS-1 cells were kindly provided by Y. Obata. Jay Unkeless and Andrew Luster generously donated mouse S49.1 and J774 and rat RBL-1 cell lines. HeLa cells were the gift of Selina Chen-Kiang. Elizabeth Lacy kindly provided mouse brains.

Preparation of nuclear extracts. All procedures were performed at 4°C, and all buffers contained 0.5 mM dithiothreitol and 0.5 mM phenylmethylsulfonyl fluoride. Nuclei were prepared as described previously (11), except that 0.5% Nonidet P-40 was used to enhance cellular lysis. Nuclear extracts were prepared either by elution of MEL cell nuclei with 0.42 M KCl (49) or by Dounce homogenization in 0.42 M KCl as described by Dignam et al. (11). Extracts prepared by both methods were dialyzed into buffer containing 20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; pH 7.9), 20% glycerol, 0.2 mM EDTA, and 100 mM KCl, and any precipitates were removed by centrifugation at 25,000 × *g* before electrophoretic assay or column chromatography.

Chromatography of DNA binding proteins. Phosphocellulose (P11; Whatman, Inc.), DEAE-cellulose (DE52; Whatman), hydroxyapatite (HAP; Bio-Gel HTP; Bio-Rad Laboratories), Ultrogel AcA34 (LKB Instruments, Inc.), and heparin-agarose (HepAg; Bio-Rad) were prepared according to the instructions of the manufacturer. Routinely, 1 × 10^{10} to 2 × 10^{10} MEL cells were processed to give 50 ml of nuclear extract at a protein concentration of 3 mg/ml. This material was passed first over a 12-ml phosphocellulose column. The flowthrough was collected and passed over a 10-ml DE52 column. Protein in the flowthrough fraction was loaded onto a 6-ml HAP column, and the active fraction was eluted with five column volumes of a buffer containing 10 mM KPO_4 (pH 7.5), 20% glycerol, 0.2 mM EDTA, and 200 mM KCl. The 200 mM KCl eluate was then precipitated by the addition of 0.106 g of $(NH_4)_2SO_4$ (20% of saturation) per ml. This precipitate was discarded, and the supernatant was further precipitated by the addition of 0.113 g of $(NH_4)_2SO_4$ (40% of saturation) per ml. This precipitate was dissolved in 1.5 ml of buffer containing 20 mM HEPES (pH 7.9), 20% glycerol, and 0.2 mM EDTA and chromatographed on a 80-ml (1.5- by 40-cm) Ultrogel AcA34 column. The binding activity of individual fractions was determined, and peak fractions were pooled. The KCl concentration was adjusted to 50 mM, and the sample was loaded onto a 2-ml heparin-agarose column equilibrated with a buffer containing 50 mM KCl. The column was then washed with 10 ml of buffer containing 75 mM KCl. This step removed a large number of contaminating polypeptides and sacrificed a small amount of binding activity specific for the α_1 promoter region (designated αA). The bulk of αA binding activity was then eluted with 10 ml of buffer containing 200 mM KCl. Each fraction was individually assayed for binding activity, and peak fractions were pooled. Nuclear extracts and column chromatographic fractions were either used immediately or were frozen in liquid nitrogen and stored at −70°C.

Electrophoretic mobility shift binding assay. The cloned fragments spanning the α_1-and β_{maj}-globin genes which were used in these studies have been described previously (8). The albumin cDNA clone was a gift from Shirley Tilghman (30). Allen Oliff kindly provided a clone containing the U3 region of the Friend murine leukemia virus long terminal repeat (31). DNA inserts were prepared by digestion with the appropriate restriction endonucleases and electroelution, and the digested ends were filled with [^{32}P]dXTPs using the large fragment of DNA polymerase I. The procedure of Strauss and Varshavsky (49) was used to perform the binding reactions. Briefly, 1 ng of ^{32}P-labeled αA DNA, the desired amount of *Escherichia coli* competitor DNA, and 5 μl of nuclear extract were mixed together in a final volume of 25 μl in 0.1% Triton X-100–4% glycerol–1 mM EDTA–10 mM 2-mercaptoethanol–10 mM Tris hydrochloride (pH 7.4). Nuclear extracts and fractions from the phosphocellulose, DE52, and HAP columns were first adjusted to a protein concentration of 200 μg/ml before assay. Column fractions subsequent to chromatography on HAP, with protein concentrations being less than 200 μg/ml, were used directly. The protein concentration was determined by the method of Bradford (2). After incubation for 30 min at room temperature, samples were loaded onto low-ionic strength, 4% polyacrylamide gels and run at 15 V/cm for 3 to 6 h at 4°C. After electrophoresis gels were dried and subjected to autoradiography overnight at −70°C with intensifying screens.

Identification of fragments of αA containing the factor binding site. Fragment αA was purified from a 5% polyacrylamide gel as described previously (34, 48) and digested with the indicated restriction endonucleases (see below). Digests were subjected to electrophoresis on 5 or 7% polyacrylamide gels, and appropriate fragments were electroeluted. Eluted fragments were end labeled in a Klenow reaction and tested for factor binding by the mobility shift assay.

DNase I and dimethylsulfate protection mapping. The 630-base-pair (bp) αA fragment was end labeled either by using the large fragment of *E. coli* DNA polymerase I (Klenow frament; New England Biolabs) or by treating the fragment with calf intestinal phosphatase (Boehringer Mannheim Biochemicals) followed by labeling with T4 polynucleotide kinase (New England Biolabs). After organic extractions and ethanol precipitation, DNA was digested with *Msp*I (New England Biolabs). A 190-bp fragment containing the α_1-globin CCAAT box region was recovered from a 5% polyacrylamide gel (34, 48). Recovered DNA was resuspended at 1 ng/μl in 10 mM Tris (pH 7.4)–1 mM EDTA. Reaction mixtures were prepared as described for the electrophoretic binding assay. All reactions contained 10 ng of unlabeled *E. coli* DNA. For DNase I protection studies, after 30 min at room temperature, $MgCl_2$ was added to 5 mM, and samples were treated with 50 ng of DNase I (Worthington Diagnostics) for 1 min at room temperature. Digestion was stopped, and mixtures were prepared for sequencing gels as described by Bram and Kornberg (3). For dimethyl sulfate protection experiments, samples were treated with 1 μl of dimethyl sulfate for 2 min at 20°C. Methylation was stopped and samples were precipitated as described previously (37). Samples were subsequently treated either to cleave at A and

G (36) or only at G (37). Parallel sequencing reactions were prepared as described previously (37).

Construction of deletion mutants of αA. Fragment αA was removed from a pUC9 vector by digestion with *Hin*dIII and *Eco*RI. The purified fragment was treated with nuclease *Bal*31 (New England Biolabs; 1 U/8 μg of the 630-bp αA DNA at 30°C). Samples were removed every 2 min, and the reaction was stopped by the addition of ethylene glycol-bis-(β-aminoethyl ether)-*N,N, N',N'*-tetraacetic acid (EGTA) to 20 mM. A portion of each sample was analyzed by 5% polyacrylamide gel electrophoresis with a Tris-borate buffer system (34) to evaluate the extent of *Bal*31 digestion. DNA termini were made blunt for ligation in an end-filling reaction using the Klenow fragment of *E. coli* DNA polymerase I (1 U/μg of DNA). Deleted promoter fragments were cloned into the *Sma*I site of plasmid pUC13. Miniature preparations of insert-containing plasmids were treated with *Eco*RI and *Hin*dIII and analyzed on 5% polyacrylamide gels. Deletions were ranked by size, end labeled, and characterized for factor binding by the electrophoretic mobility shift assay. Deletion endpoints were determined by DNA sequencing using commercially available plasmid sequencing primers (New England Biolabs) and the dideoxy sequencing procedure (43).

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Fractions containing 10 μg of protein from the DE52 flowthrough and Ultrogel peak fractions were precipitated with acetone (24), washed with ether, dried, and taken up in sodium dodecyl sulfate (SDS) sample buffer. Fractions from a HepAg column (flowthrough, 75 and 200 mM eluates) were also concentrated by acetone precipitation, except that 10 μg of linear acrylamide (49) per ml was added to increase the efficiency of protein precipitation because of the low protein concentrations in the latter stages of factor enrichment. Precipitated samples were then separated by electrophoresis on polyacrylamide slab gels (8%) which were run using the buffer conditions of Laemmli (32). Gels were stained for protein using silver (ICN Pharmaceuticals Inc.). We noted on multiple occasions that several of the proteins, including the species of 64 to 68 kilodaltons (kDa) did not stain well with silver. Counterstaining with Coomassie blue or the use of other commercially available silver-staining reagents were not effective in enhancing the stained patterns of the 64- to 68-kDa polypeptides. Relative molecular weights were determined by reference to commercially obtained molecular weight markers (Bethesda Research Laboratories).

RESULTS

Enrichment of a factor that binds the α-globin gene promoter. To identify and purify proteins which might account for some of the observed differences in murine α- and β-globin gene expression, we performed experiments to detect specific interactions of nuclear proteins with globin gene DNA fragments. An electrophoretic mobility shift assay (17, 49) was used to resolve specific protein-DNA interactions from nonspecifically bound and free DNA. In this assay ^{32}P-labeled DNA fragments were incubated with nuclear extracts or column chromatography fractions 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 incubation were visualized as discrete bands with reduced mobilities when compared with free DNA. Nonspecific interactions were visualized as a smear (Fig. 1).

Initially, salt extracts of MEL cell nuclei were analyzed for proteins which stably bound any of several cloned fragments spanning the murine α- and β-globin genes (Fig. 1A). A preliminary screen showed that one 630-bp fragment of the α_1-globin gene (αA) formed a stable complex with factors in crude nuclear extracts (Fig. 1B). Treatment of extracts with proteinase K or SDS abolished activity, and the heating of extracts to 100°C for 5 min decreased activity by more than 90% (data not shown). We then used the mobility shift assay to fractionate the active protein factor from total nuclear extracts (subsequently prepared by the method of Dignam et al. [11]) by ion-exchange, gel filtration, and affinity chromatography (Fig. 1).

Binding activity for fragment αA was contained exclusively in flowthrough fractions of phosphocellulose (P11; P11 FT) and DEAE (DE52; DE52 FT) ion-exchange resins using a buffer system containing 100 mM KCl (Fig. 1C and D; for convenience, only peak activity fractions are displayed). Pooled P11 and DE52 flowthrough fractions were applied to HAP, and the column was eluted with 200 mM KCl and 10 mM $KHPO_4$ to recover binding activity (0.2 M HAP; Fig. 1E). Analysis of material contained in the HAP column flowthrough revealed no αA-specific binding activity, nor was such activity found in fractions eluted from the P11, DE52, or HAP columns with buffers containing higher salt or phosphate concentrations. The 0.2 M HAP fraction was serially precipitated with ammonium sulfate to 20% (discarded) and 40% of saturation. Both the 20% ammonium sulfate precipitate and the supernatant remaining after 40% ammonium sulfate precipitation contained no αA-specific binding activity. The precipitate formed at 40% saturation of ammonium sulfate, which contained all binding activity, was applied to an AcA34 Ultrogel column. Analysis of individual Ultrogel column fractions revealed a single symmetric peak of αA binding activity. For convenience only peak activity fractions from the Ultrogel column are displayed in Fig. 1F. αA-specific binding activity eluted from the gel filtration column with the same k_{av} as a bovine serum albumin standard. Peak activity fractions from the Ultrogel column were pooled and applied to HepAg in 50 mM KCl. The HepAg flowthrough contained greater than 95% of the applied protein but no binding activity. The column was washed with 75 mM KCl (which elutes most of the remaining bound protein but less than 1% of the binding activity), and peak activity was eluted with 200 mM KCl (0.2 M HepAg; Fig. 1G). At this last stage of purification the mobility shift assay showed an additional band with a mobility less than that of the major species of complexed DNA (Fig. 1G). The exact nature of these complexes is unknown, but similar patterns have been observed by others by using homogeneous or highly enriched factor preparations (17, 49). One explanation for the extra bands in the mobility shift assay is that after the final chromatographic step, binding activity is sufficiently enriched to occupy lower affinity sites within the α-globin promoter region. Alternatively, a second DNA binding protein might be emerging. DNase I and dimethyl sulfate protection patterns generated by the Ultrogel and 0.2 M HepAg fractions, however, failed to reveal evidence of a second protein or a second binding site (see below).

The binding activity which we enriched from MEL cells was neither mouse nor erythroid specific. We were able to enrich a similar binding activity from a variety of cultured cell lines. These included mouse T (S49.1), B (NS-1), and macrophage-like (J774) cell lines, as well as HeLa cells and a rat basophilic leukemia cell line (RBL-1). A similar factor was also present in nuclear extracts of mouse brain (data not shown). In addition, we noted no gross change in the level of

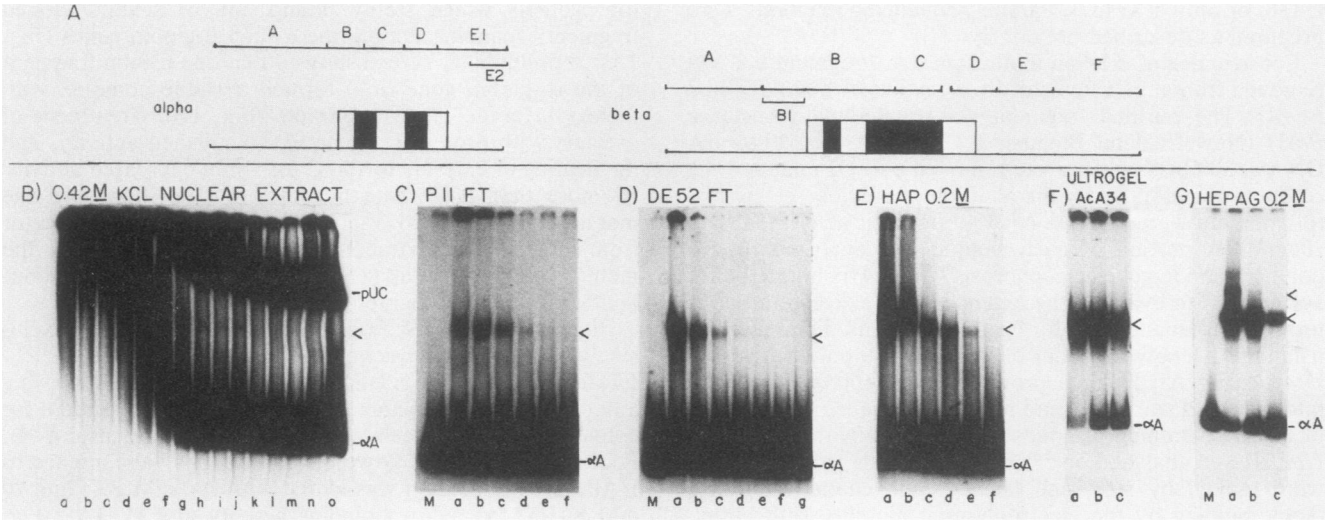

FIG. 1. Use of an electrophoretic assay to detect and monitor purification of a nuclear protein which binds to the promoter region of the α_1-globin gene. (A) The α_1- and β_{maj}-globin genes are represented at the top of the figure, beginning at the cap site and ending at the poly(A) addition site. Open boxes represent exons and filled boxes represent introns. Cloned fragments that span the gene and that were used in this study are indicated by letter designations. The lengths of the various fragments are indicated in the text and have been described previously (8, 46). (B) A total of 1 ng of ^{32}P-labeled DNA was mixed with 1 μg of 0.42 M KCl extract from uninduced MEL cell nuclei and increasing amounts of unlabeled *E. coli* competitor DNA (sonicated to an average length of 1.5 kilobase). After a 30-min incubation at room temperature, samples were subjected to electrophoresis on low-ionic-strength 4% polyacrylamide gels at 4°C, as described previously (49). Lanes a through o show DNA incubated in the presence of nuclear extract together with 0, 10, 25, 50, 100, 250, 500, 750, 1,000, 2,500, 5,000, 10,000, 25,000, 50,000, and 100,000 ng of *E. coli* DNA, respectively. The bound fragment is designated by an arrowhead, and unbound αA DNA is also indicated. In this experiment the plasmid vehicle (indicated as pUC) containing αA was digested with restriction endonucleases *Hin*dIII and *Eco*RI, and both plasmid and insert were labeled by filling in digested ends with the large fragment of *E. coli* DNA polymerase I in the presence of ^{32}P-labeled nucleotide triphosphates. In all subsequent reactions (panels C through G) the αA DNA insert was purified from the plasmid by restriction endonuclease digestion and electroelution from polyacrylamide gels (48). Crude nuclear extracts were sequentially subjected to ion-exchange chromatography, differential ammonium sulfate precipitations, and exclusion and affinity chromatography to generate protein fractions enriched for αA DNA binding activity (see text). At each stage the electrophoretic assay was used to monitor the extent of αA binding activity. The active protein fractions were found in the flowthrough of phosphocellulose (P11 FT; panel C) and DE52 (DE52 FT; panel D) columns, and in the 200 mM KCl 10 mM KPO_4 eluate of HAP columns (HAP 0.2; panel E). Peak αA binding activity eluted from an AcA34 Ultrogel column with a k_{av} equal to that of bovine serum albumin (panel F). Further chromatography on HepAg gave a peak fraction which eluted in 200 mM KCl (HepAg 0.2; panel G). In panels C through G, the arrowhead indicates the bound fragment, and the lane designated M shows the autoradiographic pattern of 1 ng of αA DNA (5,000 cpm/ng) run in the absence of added extract. Lanes indicated by letters contain the amounts of competitor DNA described for panel B.

factor present in uninduced MEL cells compared with the level in MEL cells cultured in the presence of inducer for 48 h. Factor yields, however, were substantially reduced if cells were growing poorly (data not shown).

The material from Ultrogel or HepAg columns bound 70 to 90% of the input ^{32}P-labeled αA DNA either in the absence of any competitor or in the presence of 400 ng of *E. coli* DNA per ml. By comparison, less than 1% of labeled input DNA

TABLE 1. Enrichment protocol

Fraction	Vol (ml)	Amt of protein			Sp act[a]	Total μg of protein per reaction	Total activity[b]	
		mg/ml	total mg	% total protein			U (×10^{-6})	% recovery
Salt extract[c]	57.5	3.9	224.25	100	29.5	1	6.6	100
Phosphocellulose flowthrough	115	0.96	100.40	49.2	68	1	7.5	114
DE52 flowthrough	185	0.33	60.10	26.8	200	1	12	182
0.2 M HAP	43.5	0.7	30.45	13.6	269	1	8.2	124
0.2 M HAP, 40% $(NH_4)_2SO_4$ precipitate	2.1	5.2	10.92	4.9	337	1	3.7	56
Ultrogel AcA34 peak	8	0.130	1.04	0.46	1,491	0.65	1.6	24
0.2 M HepAg peak	6	0.001	0.006	0.003	46,600	0.005	0.28	4

[a] Binding reactions were performed with protein samples from each stage of the enrichment scheme, as illustrated in Fig. 1. Conditions which gave maximal binding were established independently for each enrichment step. Binding reactions were performed under optimal conditions using an αA fragment labeled to a specific activity of 1,400 cpm/ng. After binding, bound and unbound DNA was separated by electrophoresis on low-ionic-strength polyacrylamide gels and visualized by autoradiography, and appropriate bands were cut out and counted by liquid scintillation. The counts per minute of input DNA in the bound complex was used as a measure of the binding activity in that fraction. The specific activity of factor binding is calculated as the counts per minute of total input DNA resolved as a bound complex per microgram of input protein under maximal binding conditions.

[b] Units of total activity are calculated as specific activity multiplied by the total micrograms of protein at each step.

[c] Starting material, 2 × 10^{10} cells (as described by Dignam et al. [11]).

was bound in the presence of optimal amounts of competitor DNA by using crude nuclear extracts (Fig. 1F or G versus B). To confirm the impression that the purification scheme results in a substantial enrichment of αA-specific binding activity, the mobility shift assay was used to quantitate factor enrichment. In experiments comparable to those shown in Fig. 1 the amount of probe bound per microgram of input protein was determined by cutting out and counting the radiolabeled αA DNA in a complex at each purification step (Table 1). These results show that binding activity is enriched 1,600-fold, with 4% of the total activity being recovered by our protocol. The recovery of total activity was approximately 60% until the last two chromatographic steps on Ultrogel and HepAg columns (Table 1). During these two steps the protein concentration dropped three orders of magnitude, and this decrease might have contributed to the loss of activity. After HepAg chromatography, for example, very low protein concentrations were obtained (less than 1 μg/ml), and the binding activity after this step was highly labile. The addition of inert carrier protein such as ovalbumin (50 μg/ml) retarded decay of binding activity, but other agents such as glycerol (50%), detergents (Nonidet P-40 or Triton X-100), or $MgCl_2$ did not stabilize binding activity, either alone or in combination.

The data in Table 1 provide a useful quantitative estimate of factor enrichment. These estimates, however, are subject to a potential source of error. The bands seen in the electrophoretic assay are the final product of multiple competing reactions between nuclear proteins and DNA. The relative concentrations of competing protein species are likely to change unpredictably as proteins are eliminated or enriched during the various chromatographic stages. The elimination of competing species, for example, probably accounts for calculated recoveries exceeding 100% after the P11, DE52, and HAP steps (Table 1).

Enriched factor binds to the α_1-globin gene CCAAT box. To determine which sequences within the α-globin promoter were responsible for factor binding in vitro, we first performed mobility shift assays on several restriction endonuclease fragments of the 630-bp α-globin gene promoter region (Fig. 2). Individual fragments were incubated with factor preparations, and binding was determined by the mobility shift assay. The results of these experiments show that factor binding does not require any sequences upstream (to the left) of a *Hae*III site at −132 (Fig. 2). The factor biding site maps exclusively to the right-hand end of the fragment and is completely contained within a 77-bp *Hae*III-*Bst*NI fragment. This fragment, which extends from −55 to −132, contains the α-globin gene CCAAT box but does not include the TATAA box (Fig. 2).

To characterize the factor binding site in more detail, we performed DNase I and dimethyl sulfate protection experiments (18, 47). For these experiments we digested the 630-bp αA fragment with *Msp*I to give fragments of 190 and 440 bp. Only the 190-bp fragment interacted with factor preparations using the mobility shift assay (Fig. 2). Figure 3A shows that both the Ultrogel and 0.2 M HepAg fractions contain a factor which specifically protects 26 to 28 bp of DNA centered on the α_1-globin gene CCAAT box from digestion by DNase I. No other regions on the *Msp*I fragment were protected from DNase I digestion (data not shown). Dimethyl sulfate protection studies (Fig. 3B) showed that the two Gs (pairing with the Cs) and two As that comprise the first four nucleotides of the CCAAT box are protected from methylation by factor binding. These are the only bases that show strong protection from dimethyl sulfate, although other nearby

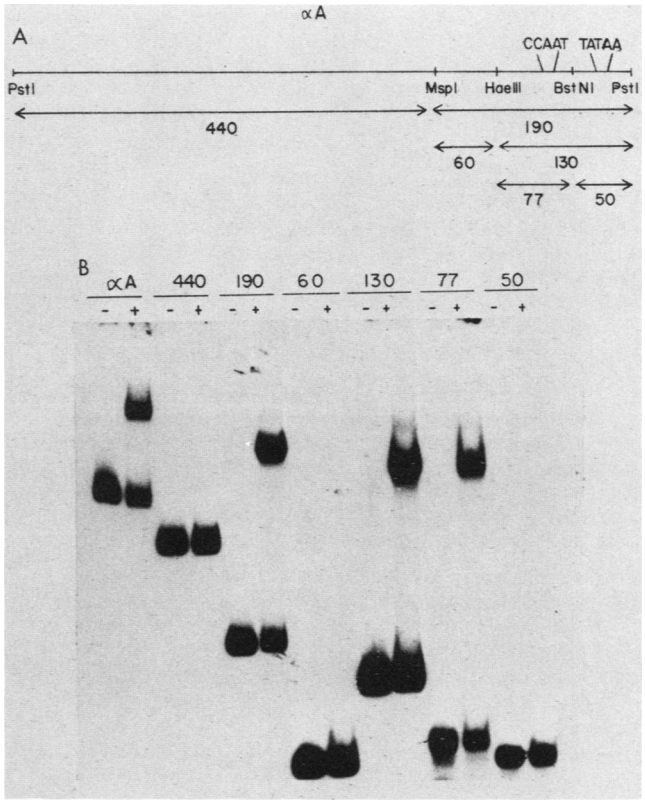

FIG. 2. A 77-bp fragment containing the α-globin gene CCAAT but not TATAA box contains all sequences required for factor binding. Fragment αA is diagrammed in panel A; relevant restriction endonuclease sites and the relative location of the CCAAT and TATAA boxes are shown. The origins of various subfragments obtained by incubating αA with the restriction endonucleases shown are indicated numerically according to their lengths (in base pairs). (B) Mobility shift assays conducted on the individual fragments indicated above. Fragments other than αA are indicated by their lengths (in base pairs). Each fragment was end labeled and incubated either in the absence (−) or presence (+) of enriched factor. Reactions were analyzed for factor binding on low-ionic-strength polyacrylamide gels, as described in the legend to Fig. 1. Fragments labeled αA, 190, 130, and 77 showed bands with shifted mobilities when incubated with factor. No factor binding to fragments labeled 440, 60 or 50 was detected.

bases within the DNase I-protected region were occasionally weakly protected. These results (Fig. 3C) show that a factor present in the Ultrogel and HepAg fractions interacts tightly with the CCAAT box of the α_1-globin gene.

Attempts at performing DNase I and dimethyl sulfate protection experiments with the fractions generated early on in the purification scheme (phosphocellulose flowthrough, DE52 flowthrough, 0.2 M HAP) did not produce useful results. The mobility shift data in Fig. 1 and Table 1 suggest a reasonable explanation for this result. Fractions from early stages of factor enrichment have a low specific binding activity. As a consequence less than 50% of the input DNA was found in a specific complex. Because the binding site was not saturated, protected sites could not be recognized against the much larger background of DNase I cuts in unbound DNA. Thus, the extent of DNA binding by the enriched factor, as assayed by mobility shift, correlated well with the ability to footprint the template.

The preceding experiments showed that a 77-bp *Hae*III-*Bst*NI fragment containing the α-globin CCAAT (but not

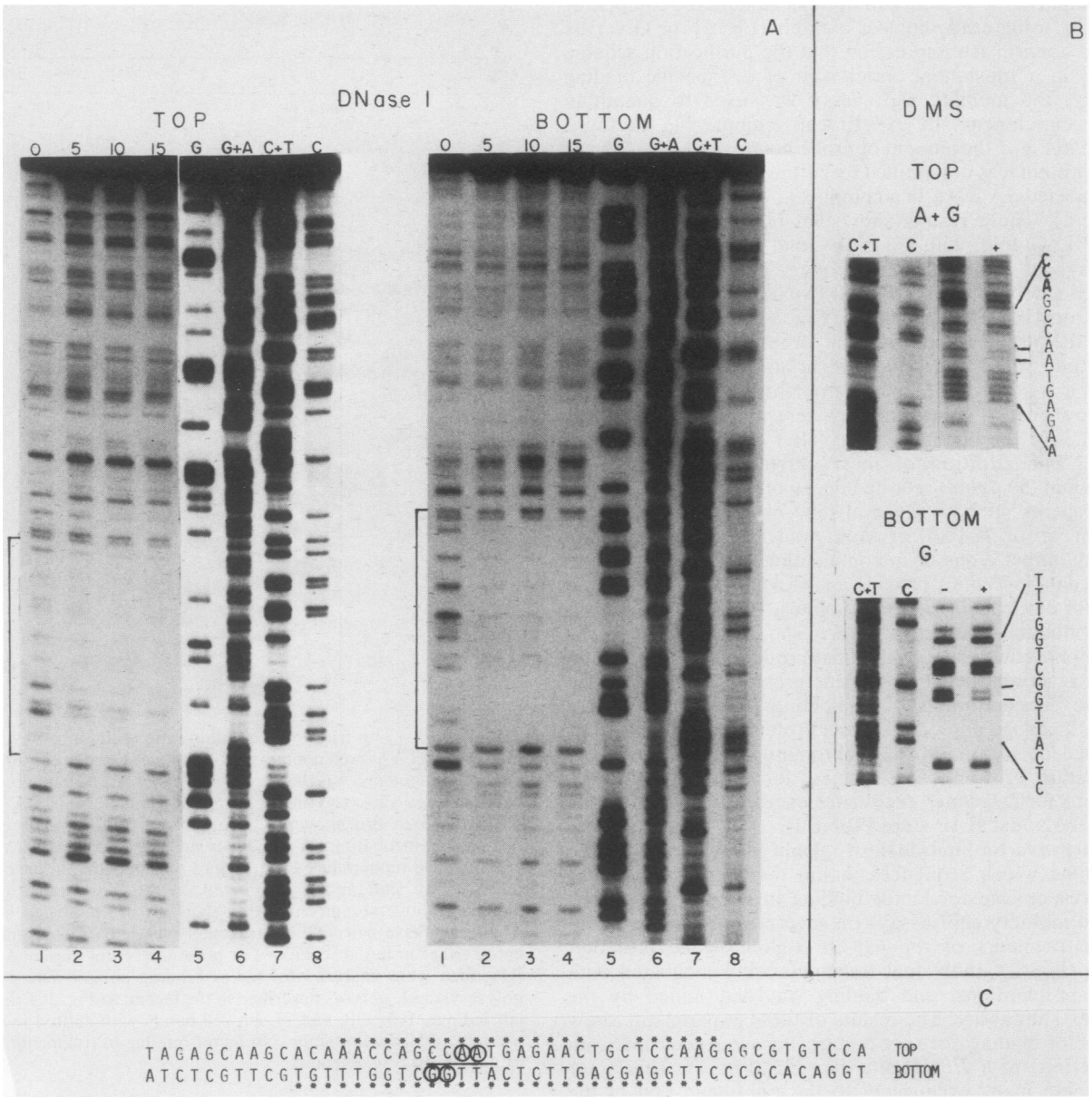

FIG. 3. DNase I footprint and dimethyl sulfate protection analysis of the interaction of the binding protein with the 5′ end of the α_1-globin gene. The top or bottom strand (see panel C) of a 190-bp fragment containing the α_1-globin gene CCAAT box was end labeled, as described in the text. Samples of DNA (1 ng) were incubated either without (lane 1) or with (lanes 2 through 4) increasing amounts of an enriched nuclear fraction, as described in the text. The volume of extract (in microliters) added to each reaction is indicated above each lane. (A) Top strand: Samples were incubated with enriched nuclear extract from an Ultrogel column. Bottom strand: Samples were incubated with enriched factor from a HepAg column. Material from Ultrogel or HepAg columns gave identical results with both strands. The base modification reactions used in parallel sequencing reactions (lanes 5 through 8) are indicated at the top of the autoradiographs. The bracket at the left of each autoradiograph indicates the region protected from cleavage by DNase I. (B) Samples were incubated without (−) or with (+) 15 μl of nuclear extract from an Ultrogel column. Samples were exposed to dimethyl sulfate (DMS) and prepared for strand cleavage at A and G (top strand) or G alone (bottom strand), as described in the text. Bases showing resistance to methylation are indicated by vertical lines to the right of the autoradiographs. A short stretch of DNA sequence around the protected region is also indicated. Only the region protected from methylation is indicated. No other sequences showed either protection or enhancement of base methylation. Parallel sequencing reactions (C+T and C reactions only) are also indicated. (C) Summary of bases near the α_1-globin gene CCAAT box protected from digestion by DNase I (indicated by the asterisks) or methylation by dimethyl sulfate (indicated by the circled bases). The CCAAT box is underlined, and the two strands designated top and bottom are also indicated.

TATAA) box contains all sequences required for factor binding in the mobility shift assay (Fig. 2). A second assay, DNase I footprinting, showed that enriched preparations contain a factor that interacts with 26 to 28 bp within the *Hae*III-*Bst*NI fragment centered on the CCAAT box. To determine whether these two assays indeed reflect factor binding to the same site, we generated a series of small deletions in αA using the exonuclease *Bal*31. Deleted fragments were cloned, and clones were ranked by their insert size. After inserts were radiolabeled, they were tested for

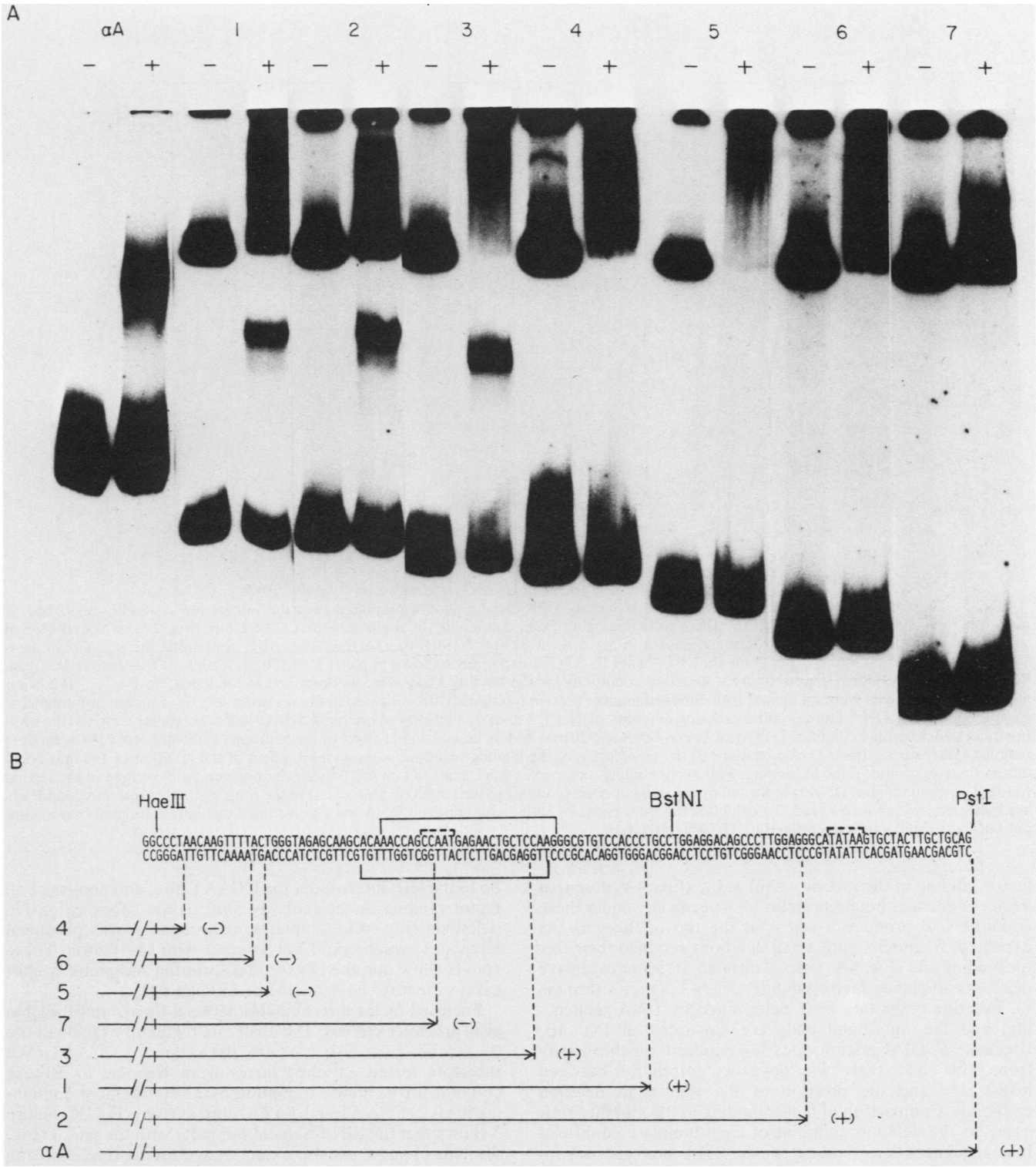

FIG. 4. Deletion analysis of the factor binding site. Fragment αA was treated with nuclease *Bal*31. Clones containing αA deletions were prepared, ranked by size, and tested for factor binding (see text). (A) Mobility shift assays were performed with purified fragment αA or one of several deletions (indicated numerically at the top of the panel). Fragments were incubated in the absence (−) or presence (+) of enriched factor. Binding was assayed as described in the legend to Fig. 1. Unlike fragment αA, the various deletions were not purified from vector sequences, which are seen at the top of the autoradiogram. Vector sequences bind nuclear proteins to produce a smear under these conditions. Factor binding to deletions 1 through 3 is visualized by the appearance of a band, the mobility of which is between that of the insert and the vector. Deletions 4 through 7 contain no factor binding site. (B) Deletion endpoints for fragments 1 through 7 were determined (see text), and ends near the known binding site (Fig. 2) are indicated. Relevant restriction endonuclease sites are also shown. The *Pst*I site marks the end of αA. All sequences required for factor binding are contained on a *Hae*III-*Bst*NI fragment (Fig.2). The region protected in the footprinting assay (Fig. 3) is indicated, as are the CCAAT and TATAA boxes. Deletion endpoints are shown for fragments which bind (+) or do not bind (−) factor in the mobility shift assay shown in panel A.

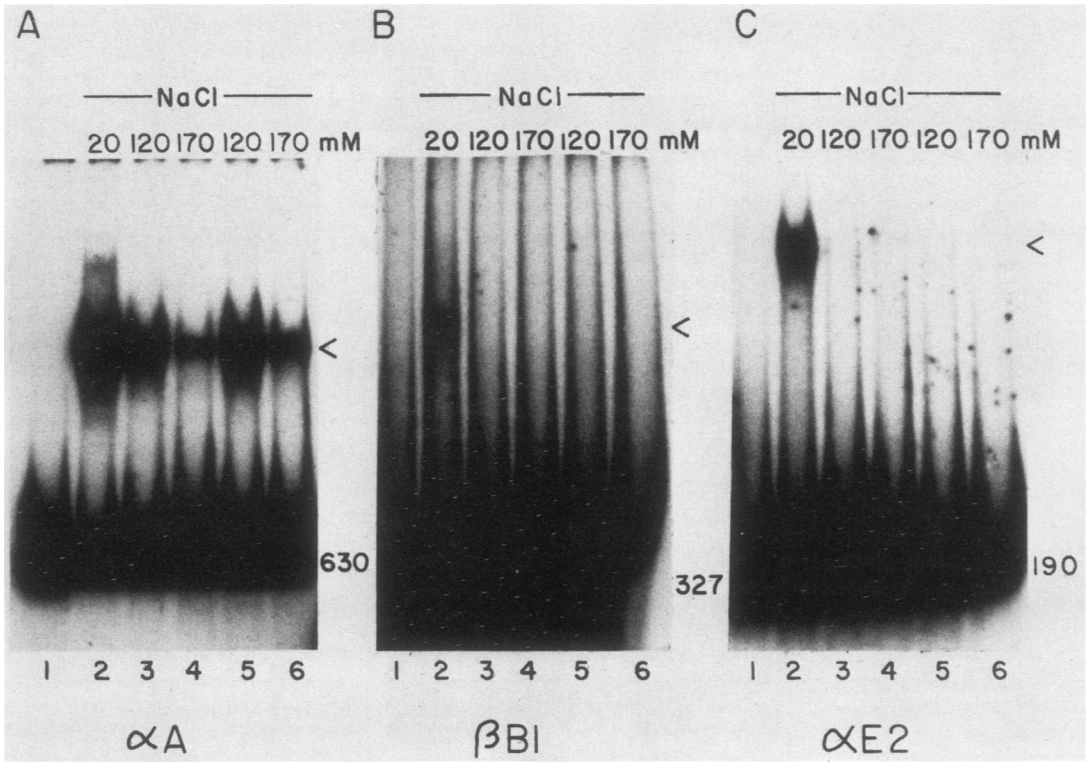

FIG. 5. The CCAAT box binding protein interacts weakly with the β_{maj}-globin promoter, and the interaction is readily abolished at moderate ionic strengths. The 630-bp *Hin*dIII-*Eco*RI fragment of αA, containing the α-globin gene CCAAT box (Fig. 1A), was end labeled with [^{32}P]dXTPs and used in the experiment shown in panel A. A 327-bp *Hin*dIII-*Sau*3A fragment, βB1, containing the β_{maj}-globin gene CCAAT box was used in the experiment shown in panel B. For the experiment shown in panel C, a 190-bp *Alu*I-*Sac*I fragment, αE2, from the 3′ end of the α_1-globin gene was used. Reaction conditions for the binding assay were as described in the legend to Fig. 1. ^{32}P-labeled globin DNA fragments were incubated with enriched nuclear protein fractions (Ultrogel peak) in the presence of a fixed amount of unlabeled *E. coli* competitor DNA (10 ng) and increasing amounts of NaCl. Lanes 1, Patterns of unbound DNA; lanes 2, binding activity for each fragment under optimal conditions (10 ng of *E. coli* DNA and 20 mM NaCl); lanes 3 and 5, effect of the addition of 100 mM NaCl to the binding reaction either during (lane 3) or after (lane 5) the completion of the binding reaction; lanes 4 and 6, effect of the addition of 150 mM NaCl during (lane 4) or after (lane 6) binding. Following binding, complexes were analyzed on 4% polyacrylamide gels as described in the legend to Fig. 1. Unbound input DNA is indicated by the number corresponding to the length (in base pairs) of the fragment. DNA-protein complexes are indicated with an arrowhead. To calculate the ratio between bound and unbound DNA for a particular fragment, individual bands were cut out from dried gels and counted by liquid scintillation.

factor binding in the mobility shift assay (Fig. 4A; note that vector sequences bound proteins nonspecifically under these conditions to produce a smear at the top of the gel). As expected, fragments with small deletions retained their factor binding site (Fig. 4A, lanes 1 through 3). More extensive deletions abolished factor binding (Fig. 4A, lanes 4 through 7). Deletion endpoints were determined by DNA sequencing, and the right-hand ends are illustrated in Fig. 4B. (Because *Bal*31 digestion does not proceed synchronously from both ends, there was no exact correlation between insert size and the position of the right-hand deletion endpoint). Comparison of factor binding in the mobility shift assay to the deletion endpoint of each template confirmed that sequences surrounding the TATAA box are not involved in factor binding and strongly suggested that the DNase I footprint region corresponds to the factor binding site. Three deletions (Fig. 4B, templates 1 through 3) removed the TATAA element but still bound factor in the mobility shift assay. One of these deletions (Fig. 4B, template 3) deleted four bases that were protected from digestion by DNase I in wild-type templates. This result suggests that the minimum factor binding site is somewhat smaller than the 26- to 28-bp region identified by footprinting. The next deletion (Fig. 4B, template 7) extended an additional 14 bp to the left, interrupted the CCAAT box, and abolished all factor binding in the mobility shift assay. More extensive deletions (Fig. 4B, templates 4 through 6) also produced templates which could not interact with the factor. These results show that the DNase I footprinting and mobility shift assays identify the same factor binding site.

Enriched factor discriminates between the α- and β-globin gene promoter regions. The upstream regulatory region of the β_{maj}-globin gene also contains the sequence CCAAT. We therefore tested enriched factor from Ultrogel or HepAg columns for its ability to bind to a 327-bp fragment containing the β_{maj} CCAAT and TATAA boxes (βB1; Fig. 1). Figure 5 shows that the βB1 fragment interacts with the factor (Fig. 5B, lane 2) but to a lesser extent than does αA (Fig. 5A, lane 2). Indeed, it is possible to abolish binding to βB1 by the addition of NaCl to 120 or 170 mM either during (Fig. 5B, lanes 3 and 4) or after (Fig. 5B, lanes 5 and 6) the binding reaction. Identical treatments reduced, but did not eliminate, complexes formed with αA DNA (Fig. 5A, lanes 3 through 6). This result suggests that sequences flanking the CCAAT core sequence of the factor binding site on the α_1-globin gene promoter have important roles in stabilizing the complex. Comparison of the nucleotide sequence of the αA (5′-CAGCCAATGAG-3′) and βB1 (5′-AGGCCAATTCT-3′)

CCAAT box regions shows that they share only 1 of 6 bases outside of the CCAAT box itself, although 6 of 11 nucleotides are shared overall. Because of the small amount of DNA that is complexed when fragment βB1 is used as a probe (less than 1% of input DNA), it was not possible to perform DNase I or dimethyl sulfate protection studies under our standard conditions.

When we retested enriched factor preparations against radiolabeled globin gene fragments other than αA and βB1, the majority of cloned inserts spanning the α_1- and β_{maj}-globin genes (Fig. 1A) again failed to give specific protein-DNA complexes in the electrophoretic assay. Three subclones, however, demonstrated weak binding. These were two internal fragments (αB and βD; data not shown), and a fragment containing a 3'-flanking region (αE2; Fig. 5C). Sequence analysis of these fragments showed small regions of homology to the nucleotides protected in αA. For example, a sequence in αE2 (5'-CAGCCATGTG-3') shared 9 of 11 bases with αA (allowing for a 1-bp gap). Presumably, these limited regions of sequence homology account for the weak binding of enriched factor to the fragments. It is important, though, that the percentage of input DNA found in complexes using these fragments never exceeded 20% compared with 70 to 90% obtained with fragment αA (compare Fig. 5A through C, lanes 2). Our ability to visualize these interactions was most likely the result of a relative increase in specific factor concentration during the course of the fractionation and the low ionic strengths (20 mM) which were used in the incubation mixtures. Indeed, it was possible to abolish binding to all globin gene fragments other than αA by the addition of NaCl to 120 mM either during or after the binding reaction (Figure 5A through C, lanes 3 through 5). These results indicate that while limited amounts of binding can occur with some globin DNA sequences at low ionic strengths, specific binding stable to physiologic salt concentrations occurs only with αA. It should also be noted that the CCAAT box of the α_1-globin gene is not unique in its ability to exhibit stable factor binding at moderate ionic strengths. We have found that a DNA fragment containing the CCAAT box in the U3 region of the long terminal repeat of the Friend murine leukemia virus (31) stably binds factor in 170 mM NaCl (data not shown).

Characterization of the proteins responsible for CCAAT box binding. The molecular mass of the αA-specific binding protein was estimated to be 68 kDa by gel filtration. The data in Fig. 1G also show that at the last step in factor enrichment, approximately 70% of input DNA (1 ng of the 630-bp αA fragment) is associated with the factor. It can be calculated by using this information and two reasonable assumptions (a factor-DNA mass ratio of 1:6 and a factor-DNA binding ratio of 1:1) that 5 μl of the 0.2 M HepAg column fraction used in Fig. 1G contains 0.12 ng of factor (24 ng of factor per ml). This value is a minimum estimate because the addition of up to 50 ng of αA DNA to a single binding reaction still results in factor binding by 70% of input DNA (data not shown). The factor preparation used in Fig. 1G was contained in a total volume of 8 ml. Using the minimum estimate of factor concentration, 190 ng of αA binding activity are contained in this volume. This amount of protein should be readily detectable by silver staining of SDS-polyacrylamide gels because the detection limit of our silver staining protocol was between 1 and 10 ng per band (data not shown).

Approximately 10 μg of protein from the latter stages of factor enrichment were concentrated by acetone precipitation, separated on 8% SDS-polyacrylamide gels (32), and stained with silver (40). To visualize proteins contained in the 0.2 M HepAg fraction (protein concentration, 1 μg/ml), the entire fraction was precipitated. Precipitates containing 10 μg of the peak activity fraction from an Ultrogel column (Fig. 6, lane 3) contained several prominent species that migrated near a 68-kDa bovine serum albumin marker. Many of these proteins were subsequently detected exclusively in the HepAg flowthrough (Fig. 6, lane 4), which was a fraction with no measureable DNA binding activity (data not shown). Thus, these prominent species cannot be responsible for complex formation or CCAAT box binding. Absent from the HepAg flowthrough were a set of three polypeptides with mobilities between 64 and 68 kDa. Of these, a 64-kDa species was particularly abundant. All three polypeptides at 64 to 68 kDa were eluted from the HepAg column over a

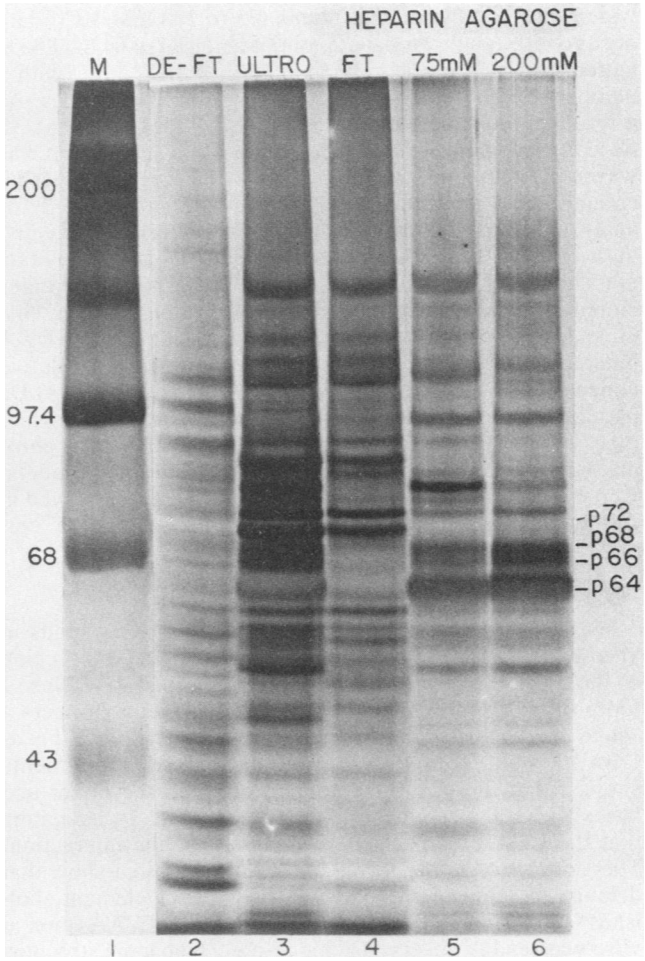

FIG. 6. SDS-polyacrylamide gel electrophoretic analysis of the polypeptide composition of column fractions. Fractions of enriched nuclear protein were precipitated with acetone, resolved by electrophoresis on 8% SDS-polyacrylamide gels, and stained with silver. Lane 1, Relative molecular mass standards ($\times 10^3$; Bethesda Research Laboratories); lane 2, protein composition of the DE52 flowthrough fraction which is representative of the starting material in the crude nuclear extract; lane 3, polypeptide content of the Ultrogel fraction containing the peak of αA DNA binding activity; lane 4, protein content of the flowthrough fraction after the Ultrogel peak material was passed over HepAg in a buffer containing 50 mM KCl; lane 5, polypeptides eluted from HepAg with 75 mM KCl; lane 6, polypeptides eluted from HepAg with 200 mM KCl. Polypeptide molecular weights were determined based on their mobilities relative to those of the protein standards in lane 1.

wide range of salt concentrations (75 to 300 mM KCl; data not shown). A substantial amount of material of 64 to 68 kDa eluted slowly from the column at 75 mM KCl, along with a number of high- and low-molecular-weight polypeptides. As a result of the slow elution at this ionic strength, the 64- to 68-kDa proteins were relatively dilute when the column was washed at this ionic strength (Fig. 6, lane 5; note that all column wash fractions were pooled and precipitated for analysis). Binding activity for αA DNA was readily detected in the 75 mM wash, but it was at a relatively low level (data not shown). The bulk of the 64- to 68-kDa polypeptides eluted at 200 mM KCl from the HepAg column (Fig. 6, lane 6), and it was this fraction which contained most of the DNA binding activity (Fig. 1G). Thus, the appearance and relative concentration of a set of three major proteins at 64 to 68 kDa parallels the fractionation of DNA binding activity. Several other minor polypeptides, however, displayed similar chromatographic behavior. To determine which polypeptide(s) were responsible for the specific binding activity observed in solution, additional analytical procedures must be employed.

DISCUSSION

We detected and enriched a protein factor that binds in vitro to a DNA fragment containing the CCAAT-box region of the α_1-globin gene promoter (Fig. 1 through 4). DNase I footprinting experiments showed that the factor protects a single binding site of 26 to 28 bp which includes the α-CCAAT box. Dimethyl sulfate protection experiments showed that there is close contact between the factor and bases within the sequence 5′-CCAAT-3′ (Fig. 3), indicating that the CCAAT box itself is essential for the interaction. This conclusion is supported by experiments that show that deletion of sequences containing the CCAAT element abolishes factor binding (Fig. 4). Binding to the α-CCAAT box in vitro was readily observed at moderately high ionic strengths (170 mM NaCl; Fig. 5), suggesting that factor binding to the α-CCAAT box could occur in vivo. No binding to a comparable region of the β_{maj}-globin promoter was observed at these ionic strengths, however (Fig. 5B). This discrimination suggests that the factor could play a role in the differential expression of the α- and β-globin genes. Examination of the polypeptides contained in our most enriched factor preparations showed a limited set of factor candidates. One protein, at 64 kDa, was particularly abundant (Fig. 6). Additional experiments are required, however, before any of the visualized proteins can be conclusively implicated in α-globin gene CCAAT box binding. Although the factor described here was isolated from MEL cell nuclei based on its ability to bind to a region of the α-globin gene promoter, we stress that it is found in a variety of cell lines and in mouse brain and therefore probably performs a general cellular function.

Comparison of the nucleotide sequence protected by the factor described here with the DNA binding domains of other nuclear factors shows that this protein is apparently unrelated to NF-1 (39), SP-1 (12), factor B (41), α protein (49), or the erythroid factor(s) specific for the chicken β-globin gene promoter (16). In addition, by chromatographic behavior and our estimation of its molecular mass (68 kDa), the protein described here does not correspond to any other previously characterized transcription factor (11, 14, 29, 42).

The strong in vitro binding of the factor to the α-globin gene CCAAT box raises the possibility that it could bind to this sequence in vivo. The CCAAT box sequence is highly conserved among mammalian globin genes, and alterations in this sequence significantly decrease globin gene expression in transient transfection assays (10, 27, 38). In addition, a similar sequence was observed upstream from most genes, identifying the CCAAT box as an important general component of the eucaryotic promoter. Although the function of the factor we identified is unknown, it seems likely that a protein which could bind to a CCAAT box or which could discriminate among different CCAAT boxes by binding in vivo only to those with preferred flanking sequences might have important global effects on gene expression. In support of this speculation, results of recent reports have shown that naturally occurring mutations in nucleotides immediately flanking the human fetal globin gene CCAAT box are associated with the inappropriate expression of fetal hemoglobin during adult life (9, 21).

In considering the possibility that binding of this factor to promoter regions in vivo could influence gene expression, one striking aspect of the results presented here is the 100-fold preference of the protein for the α- compared to the β-globin gene promoter region (Fig. 5). We noted earlier that the regulation of the α- and β-globin genes is distinctive and is likely to reflect fundamental differences in promoter function. In this regard, the high preference of the factor described here for binding to the α-globin gene CCAAT box might help to explain some aspects of the behavior of the α- and β-globin genes in transient and stable transfection studies.

Transfected genes readily avoid being packaged into a transcriptionally repressed chromatin structure (52). One consequence of the resulting open chromatin structures might be that the CCAAT boxes of transfected globin genes are accessible to generally available nuclear factors, such as the one described here. The CCAAT box of a transfected α-globin gene could then interact stably with a factor such as that described here because of its high-affinity binding site. Our results indicate that the promoter region of a transfected β-globin gene, in contrast, would constitute a low-affinity binding site for this, or a similar, factor. Hence, it is unlikely that a factor such as the one described here could influence expression of a transfected β-globin gene in any way. If this factor acted as a positive activator, expression of transfected α-globin genes would be greatly favored relative to the expression of β-globin genes in the simplified scheme outlined above. β-globin gene expression, in turn, might depend on the synthesis of erythroid-specific factors, such as those described by Emerson and co-workers (15, 16), or on linkage to an enhancer. The interaction, postulated above, between a sequence-specific DNA binding protein and a globin gene CCAAT box does not depend strongly on whether the DNA template is derived from mouse or human sources. We mention this because many of the recent transfection studies showing the distinctive behavior of the α- and β-globin gene promoters used the human sequences. It should be noted, therefore, that comparison of the DNA sequences of the mouse and human globin genes shows that the two are identical for 12 nucleotides centered on the α-CCAAT box, while the CCAAT boxes of the β-globin genes are more similar to each other than they are to α (33).

We noted earlier that stably transfected α-globin genes are expressed at a high constitutive level in the erythroid environment of MEL cells (6), which suggests that they are contained in active chromatin structures. The expression of these transfected genes is in contrast with the low level of transcription measured from the endogenous genes (46). This latter observation is consistent, however, with our earlier demonstration that the chromatin structure of the

endogenous α-globin genes is not fully active (46). These previous studies showed that in uninduced MEL cells a nuclease-hypersensitive site maps approximately 200 bp 5′ of the endogenous α_1-globin gene mRNA capping site in the absence of active α-globin gene transcription. After induction, however, additional nuclease-sensitive sites appeared. One of these sites, which was sensitive to S1 and micrococcal nuclease, maps near the mRNA capping site of the α_1-globin gene promoter and appears to include the sequences protected in vitro by the factor described here.

Evidence is accumulating that changes in the sensitivity of chromatin regions to nuclease digestion is reflected in the binding of nuclear factors in vivo (15, 16, 55). The appearance of additional nuclease-sensitive sites near the CCAAT box of the α_1-globin gene during MEL cell differentiation thus suggests that protein factors required for α_1-globin gene transcription might be binding in this region. One such protein could be the factor described here. We have no evidence as yet, however, that the site bound by this factor in vitro is occupied in vivo. Genomic sequencing studies (7, 28) will be required to determine whether the region corresponding to the DNase I footprint in vitro is also protected by the binding of nuclear factors in vivo during MEL cell differentiation.

As described above, the interaction of the factor we identified with the α-globin gene CCAAT box is postulated to have a positive regulatory effect. In this regard the role of the factor would be similar to that of the generally available transcription factor(s) hypothesized by Charnay et al. (6) in their model for globin gene regulation. Alternatively, selective binding of a protein to a CCAAT box might repress gene expression, which is a model favored by those studying the mutations, noted above, that prevent the normal developmental silencing of human fetal globin gene expression. It is possible that one class of CCAAT box might be induced and another repressed by interaction with the same protein. In vitro transcription assays will help to determine whether the factor we isolated can lead to the preferential expression or repression of α-globin gene templates.

Other mechanisms can be envisioned by which the factor described here could affect gene expression after it binds to the CCAAT box. The formation of normal nuelcosomes over essential promoter elements, for example, might be prevented by this protein which, in turn, could help to direct and accelerate binding of RNA polymerase II and other transcription factors (19). We have shown previously that the nucleosome structure of the α_1- and β_{maj}-globin genes is highly disrupted in MEL cells even before the genes are actively transcribed (8). Proteins which establish and maintain these domains have not been characterized. While the size of the α_1-globin gene binding site identified here (26 to 28 bp) makes it unlikely that the 64-kDa protein alone could directly account for the widespread (1.5 kilobase) nucleosome disruption we observed, it is possible that protein binding could indirectly cause significant alterations in chromatin structure. In vivo chromatin assembly systems (20) should allow definition of a potential role for the protein identified here in establishing a disrupted nucleosome structure over the α-globin gene.

ACKNOWLEDGMENTS

This work was supported, in part, by Public Health Service grants CA-31768 and CA-08748 from the National Cancer Institute, the Bristol-Meyers Foundation, and the Charles H. Revson Foundation. M.S. is a recipient of a Miriam and Benedict Wolf Fellowship.

LITERATURE CITED

1. **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.
2. **Bradford, M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **42**:248–254.
3. **Bram, R. J., and R. D. Kornberg.** 1985. Specific protein binding to far upstream activating sequences in polymerase II promoters. Proc. Natl. Acad. Sci. USA **82**:43–47.
4. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
5. **Chao, M. V., P. Mellon, P. Charnay, T. Maniatis, and R. Axel.** 1983. The regulated expression of β-globin genes introduced into mouse erythroleukemia cells. Cell **32**:483–493.
6. **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.
7. **Church, G. M., and W. Gilbert.** 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA **81**:1991–1995.
8. **Cohen, R. B., and M. Sheffery.** 1985. Nucleosome disruption precedes transcription and is largely limited to the transcribed domain of globin genes in murine erythroleukemia cells. J. Mol. Biol. **182**:109–129.
9. **Collins, F. S., J. E. Metherall, M. Yamakawa, J. Pan, S. Weissman, and B. Forget.** 1985. A point mutation in the ^{a}gamma-globin gene promoter in Greek hereditary persistence of fetal hemoglobin. Nature (London) **313**:325–326.
10. **Dierks, P., A. van Ooyen, M. D. Cochran, C. Dobkin, J. Reiser, and C. Weissmann.** 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit β-globin gene in mouse 3T6 cells. Cell **32**:695–706.
11. **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.
12. **Dynan, W. S., and R. Tjian.** 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. Cell **32**:669–680.
13. **Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. K. DeRiel, B. G. Forget, S. M. Weissmann, J. L. Slightom, A. E. Blechl, O. Smithies, F. E. Baralle, C. C. Shoulders, and N. J. Proudfoot.** 1980. The structure and evolution of the human β-globin gene family. Cell **21**:653–668.
14. **Egly, J. M., N. G. Miyamoto, V. Moncollin, and P. Chambon.** 1984. Is actin a transcription initiation factor for RNA polymerase B? EMBO J. **3**:2363–2371.
15. **Emerson, B. M., and G. Felsenfeld.** 1984. Specific factor conferring nuclease hypersensitivity at the 5′ end of the chicken adult β-globin gene. Proc. Natl. Acad. Sci. USA **81**:95–99.
16. **Emerson, B. M., C. D. Lewis, and G. Felsenfeld.** 1985. Interaction of specific nuclear factors with the nuclease-hypersensitive region of the chicken adult β-globin gene: nature of the binding domain. Cell **41**:21–30.
17. **Fried, M. G., and D. M. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6526.
18. **Galas, D., and A. Schmitz.** 1978. DNase I footprinting. Nucleic Acids Res. **5**:3157–3170.
19. **Gargiulo, G., F. Razvi, I. Ruberti, I. Mohr, and A. Worcel.** 1985. Chromatin specific hypersensitive sites are assembled on a Xenopus histone gene injected into Xenopus oocytes. J. Mol. Biol. **181**:333–349.
20. **Gargiulo, G., and A. Worcel.** 1983. Analysis of the chromatin assembled in germinal vesicles of Xenopus oocytes. J. Mol. Biol. **170**:699–722.
21. **Gelinas, R., B. Endlich, C. Pfeiffer, M. Yagi, and G. Stamatoyannopoulos.** 1985. G to A substitution in the distal CCAAT box of the ^{a}gamma-globin gene in Greek hereditary presistence of fetal haemoglobin. Nature (London) **313**:323–325.

22. **Green, M. R., R. Treisman, and T. Maniatis.** 1983. Transcriptional activation of cloned human β-globin genes by viral immediate-early gene products. Cell **35**:137–148.
23. **Grosveld, G. C., C. K. Shewmaker, P. Jat, and R. A. Flavell.** 1981. Localization of DNA sequences necessary for transcription of the rabbit β-globin gene in vitro. Cell **25**:215–226.
24. **Hager, D. A., and R. R. Burgess.** 1980. Elution of proteins from sodium dodeyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: results with sigma subunit of Escherichia coli RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. Anal. Biochem. **109**:76–86.
25. **Hardison, R. C.** 1983. The nucleotide sequence of the rabbit embryonic globin gene β4. J. Biol. Chem. **258**:8739–8744.
26. **Hofer, E., R. Hofer-Warbinek, and J. E. Darnell, Jr.** 1982. Globin RNA transcription: a possible termination site and demonstration of transcriptional control correlated with altered chromatin structure. Cell **29**:887–893.
27. **Humphries, R. K., T. Ley, P. Turner, A. D. Moulton, and A. Nienhuis.** 1982. Differences in human α-, β-, and δ-globin gene expression in monkey kidney cells. Cell **30**:173–183.
28. **Jackson, P. D., and G. Felsenfeld.** 1985. A method for mapping intranuclear protein-DNA interactions and its application to a nuclease hypersensitive site. Proc. Natl. Acad. Sci. USA **82**:2296–2300.
29. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell **42**:559–572.
30. **Kioussis, D., F. Faiferman, P. van de Rijn, M. B. Gorin, R. S. Ingram, and S. M. Tilghman.** 1981. The evolution of α-fetoprotein and albumin. J. Biol. Chem. **256**:1960–1967.
31. **Koch, W., W. Zimmermann, A. Oliff, and R. Friedrich.** 1984. Molecular analysis of the envelope gene and long terminal repeat of Friend mink cell focus-inducing virus: implications for the functions of these sequences. J. Virol. **49**:828–840.
32. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
33. **Liebhaber, S. A., M. J. Goossens, and Y. W. Kan.** 1980. Cloning and complete nucleotide sequence of human 5′-α-globin gene. Proc. Natl. Acad. Sci. USA **77**:7054–7058.
34. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Marks, P. A., and R. A. Rifkind.** 1972. Protein systhesis: its control in erythropoiesis. Science **175**:955–961.
36. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
37. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base specific chemical cleavages. Methods Enzymol. **65**:499–560.
38. **Mellon, P., V. Parker, Y. Gluzman, and T. Maniatis.** 1981. Identification of DNA sequences required for transcription of the human α_1-globin gene in a new SV-40 host-vector system. Cell **27**:279–288.
39. **Nagata, K., R. A. Guggenheimer, and J. Hurwitz.** 1983. Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. Proc. Natl. Acad. Sci. USA **80**:6177–6181.
40. **Oakley, B. R., D. Kirsch, and N. R. Morris.** 1980. A simplified ultrasensitive silver stain for detecting proteins in polyacrylamide gels. Anal. Biochem. **105**:361–363.
41. **Parker, C. S., and J. Topol.** 1984. A Drosophila RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. Cell **36**:357–369.
42. **Samuels, M., A. Fire, and P. A. Sharp.** 1982. Separation and characterization of factors mediating accurate transcription by RNA polymerase II. J. Biol. Chem. **257**:14419–14427.
43. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
44. **Sheffery, M., R. A. Rifkind, and P. A. Marks.** 1982. Murine erythroleukemia cell differentiation: DNase I hypersensitivity and DNA methylation near the globin genes. Proc. Natl. Acad. Sci. USA **79**:1180–1184.
45. **Sheffery, M., R. A. Rifkind, and P. A. Marks.** 1983. Hexamethylenebisacteamide-resistant murine erythroleukemia cells have altered patterns of inducer-mediated chromatin changes. Proc. Natl. Acad. Sci. USA **80**:3349–3353.
46. **Sheffery, M., R. A. Rifkind, and P. A. Marks.** 1984. Gene expression in murine erythroleukemia cells. Transcriptional control and chromatin structure of the α_1-globin gene. J. Mol. Biol. **172**:417–436.
47. **Siebenlist, U., R. Simpson, and W. Gilbert.** 1980. E. coli RNA polymerase interacts homologously with two different promoters. Cell **20**:269–281.
48. **Smith, H. O.** 1980. Recovery of DNA from gels. Methods Enzymol. **65**:371–380.
49. **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.
50. **Talkington, C. A., and P. Leder.** 1982. Rescuing the in vitro function of a globin pseudogene promoter. Nature (London) **298**:192–195.
51. **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.
52. **Weintraub, H.** 1983. A dominant role for DNA secondary structure in forming hypersensitive structures in chromatin. Cell **32**:1191–1203.
53. **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.
54. **Wright, S., A. Rosenthal, R. A. Flavell, and F. Grosveld.** 1984. DNA sequences required for regulated expression of β-globin genes in murine erythroleukemia cells. Cell **38**:265–273.
55. **Wu, C.** 1984. Activating protein factor binds in vitro to upstream control sequences in heat shock gene chromatin. Nature (London) **311**:81–84.