MOLECULAR AND CELLULAR BIOLOGY, Jan. 1988, p. 381–392
0270-7306/88/010381-12$02.00/0

Vol. 8, No. 1

Detection of Two Tissue-Specific DNA-Binding Proteins with Affinity for Sites in the Mouse β-Globin Intervening Sequence 2

DEBORAH L. GALSON AND DAVID E. HOUSMAN*

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received 16 July 1987/Accepted 16 October 1987

To identify proteins from uninduced murine erythroleukemia nuclear extracts which specifically bind to sequences from the DNase I-hypersensitive region within the mouse β-globin intervening sequence 2 (IVS2), a gel electrophoretic mobility shift assay was used. Two distinct sequence-specific binding proteins were detected. The specific binding sites for these factors were delineated by both DNase I protection footprinting and methylation interference. Factor B1 bound specifically to two homologous sites, B1-A and B1-B, approximately 100 base pairs apart within the IVS2 and on opposite strands. These two regions could interact with factor B1 independently. Factor B1 was limited to cells of hematopoietic lineages. Factor B2 bound to a site approximately 5 base pairs away from the B1-A site and was limited to cells of the erythroid lineage. The limited tissue distribution of these factors and the locations of their binding sites suggest that one or both of these factors may be involved in the formation of the tissue-specific DNase I-hypersensitive site in the IVS2 of the mouse β-globin gene.

Control of expression of the β-globin gene complex has been a major focus of research on eucaryotic regulatory mechanisms. Analysis by deletion, point mutation, and the construction of chimeric genes of the elements that control gene expression indicates that the transcriptional regulation of the β-globin gene involves a complex interplay of *cis*- and *trans*-acting elements. *cis*-Acting sequences required for efficient and accurate transcription of the mouse β-major (β^M)-globin gene in vivo have been identified within the 110 base pairs (bp) immediately upstream of the mRNA cap site. They include three highly conserved regions: a CACCC box at −90, a CCAAT box at −75, and a TATA box at −30 (10, 13, 34). The CACCC box is conserved among all adult β-like globin genes. The upstream sequences, the CACCC and CCAAT boxes, are required for efficient transcription of mouse β-globin in both HeLa and murine erythroleukemia (MEL) cells. The TATA box (at −30 in most eucaryotic genes) is necessary for the accurate initiation of transcription (6).

The sequences responsible for the capacity of the β-globin gene to respond to the inducing environment of the differentiating MEL cell have been demonstrated to reside both in the 5′-flanking DNA and within the structural gene itself (10, 11, 51). These two inducible regulatory regions have been shown to be able to act independently by the stable introduction into MEL cells of chimeric genes containing either the human β-globin promoter and the gene body of another gene (human α- or γ-globin or mouse *H2K*) or vice versa (11, 51). The presence of either the β-globin promoter or the β-globin gene body is sufficient to confer inducibility upon a hybrid gene in MEL cells. The one exception is the 5′-β/α-3′ hybrid construct in which the α-globin gene body is apparently dominant over the β-globin promoter (the transfected α-globin gene is transcribed constitutively in MEL cells). In addition, transfected β-globin genes with most of the promoter deleted remain inducible, although they are not transcribed efficiently due to the lack of the CACCC or CCAAT box or both (10, 51). The precise locations of both the 5′ flank and intragenic sequences responsible for the inducibility of the β-globin gene have not yet been identified.

trans-Acting regulatory factors and their cognate sites on DNA have been found to lie within DNase I-hypersensitive regions for the chicken β-globin gene (16, 17). Tissue-specific DNase I-hypersensitive sites have been observed in the 5′ flank (at ~−200) and within the intervening sequence 2 (IVS2) (at ~+500) of the mouse β^M-globin gene (1, 4, 25, 41) (see Fig. 1 for a diagram of the mouse β-globin gene). The latter DNase I-hypersensitive site suggests that regulatory sequences may reside in the 5′ portion of IVS2. Furthermore, LaFlamme et al. (27) transfected chimeric globin genes involving the exchange of the IVS2 between the human β- and δ-globin genes into MEL cells and demonstrated that the IVS2 of human β-globin is necessary, but not sufficient, for inducible expression of the gene in MEL cells. In transgenic mice, the human β-globin gene is regulated as a mouse adult globin gene and expressed in both fetal and adult stages (9, 29, 48) and the human γ-globin gene is regulated as an embryonic globin gene (8). Chimeric γ/β-globin genes (joined at the *Bam*HI site at the 3′ end of exon 2) in transgenic mice are expressed as both embryonic and adult β-globin genes (26), further implicating sequences downstream of exon 2 in the regulation of the β-globin gene.

An alternative approach to the mutational analyses used in previous studies to identify sequences with a regulatory function is to look for the *trans*-acting proteins that bind to DNA regions suspected of having a regulatory role and then to delineate their binding sites. This approach has the added benefit of providing the means to characterize the *trans*-acting factors as well as the sequences to which they bind. An electrophoretic mobility shift assay has been developed (18, 20, 47) to detect DNA-protein interactions that may represent interactions between *trans*-acting factors and their cognate *cis*-acting regulatory sequences. The assay is based on the altered electrophoretic mobility of DNA-protein complexes relative to unbound DNA. In view of the evidence suggesting that sequences in the IVS2 of the β-globin gene may have a regulatory role, we used the electrophoretic mobility shift assay to screen sequences from the IVS2 for

* Corresponding author.

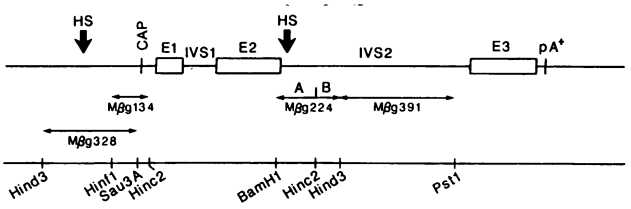

FIG. 1. Schematic representation of the murine β^M-globin gene locus. The murine β^M-globin gene is depicted with its cap (CAP) and poly(A) addition (pA^+) sites, its three exons (E1, E2, and E3), and its two intervening sequences (IVS1 and IVS2). Also indicated are the two DNase I-hypersensitive sites (HS) observed by Benezra et al. (4). The various fragments of the gene used in these studies are shown with the restriction sites used to generate them indicated underneath. The positions of the exons relative to the cap site at +1 are: E1, +52 to +144; E2, +261 to +483; and E3, +1137 to +1366.

the ability to detect sequence-specific DNA-binding proteins in nuclear extract from MEL cells. We report here the detection of two sequence-specific factors with limited tissue distributions. These factors bound to sequences within the region of the mouse β^M-globin gene IVS2 involved in the formation of a tissue-specific DNase I-hypersensitive site.

MATERIALS AND METHODS

Plasmids and DNA fragments. Plasmid pMβ1BII, kindly provided by R. Myers and T. Maniatis, contains the mouse β^M-globin gene from −106 to +2791 with a *Bgl*II linker inserted at +26 (numbering is relative to the cap site at +1). The following subclones were derived from this plasmid: 134pUC, *Cla*I$_{-106}$-*Bgl*II$_{+26}$ inserted into *Bam*HI-*Acc*I-cut pUC19; 224pUC, *Bam*HI$_{+465}$-*Hin*dIII$_{+689}$ inserted into *Bam*HI-*Hin*dIII-cut pUC19; and 391pUC, *Hin*dIII$_{+689}$-*Pst*I$_{+1080}$ inserted into *Hin*dIII-*Pst*I-cut pUC19 (Fig. 1). The inserts Mβg134, Mβg224, and Mβg391 were all excised with *Eco*RI-*Hin*dIII and then isolated from an agarose gel by electroelution. The eluted DNAs were ethanol precipitated twice, and the final DNA concentrations were determined by an ethidium bromide dot assay. Mβg224 was end labeled with Klenow fragment and [α-^{32}P]dATP by standard methods (30) and recut with *Hin*cII to generate labeled Mβg224-A and Mβg224-B, which were purified by polyacrylamide gel electrophoresis.

The two strands of an oligonucleotide designated oligo2 were synthesized by D. Lovern, Integrated Genetics, Inc., to form an *Ava*I end and an *Eco*RI end when annealed. The duplex oligo2 was used as is and was also subcloned into *Ava*I-*Eco*RI-cut pUC19, and four subclones were isolated (see Fig. 8A). Plasmid pmseHbb2, a gift from T. Glaser, contains the mouse β^M-globin gene from *Eco*RI$_{-1412}$ to *Bam*HI$_{+465}$ in pUC12. To obtain the DNA fragment Mβg328 from pmseHbb2 for use as unlabeled competitor DNA, the 808-bp *Hin*dIII$_{-344}$-*Bam*HI$_{+465}$ (Mβg808) fragment was isolated and recut with *Sau*3A to generate the 328-bp *Hin*dIII$_{-344}$-*Sau*3A$_{-16}$ fragment, which was agarose gel purified as above. To generate the labeled 113-bp *Hin*dIII$_{-344}$-*Fok*I$_{-231}$ fragment, Mβg808 was labeled with Klenow fragment and recut with *Fok*I before polyacrylamide gel purification. Plasmid pHE119, generously provided by C. Hutchinson, contains the *Bam*HI$_{+465}$-*Hin*dIII$_{+2465}$ fragment of genomic β-minor (β^m)-globin in pBR322. The DNA fragments *Bam*HI$_{+465}$-*Taq*I$_{+580}$, *Bam*HI$_{+465}$-*Hin*cII$_{+614}$, and *Taq*I$_{+580}$-*Dra*I$_{+736}$ were all used in this study. Plasmid TDp107, kindly given by T. Ley, contains the *Alu*I fragment of the wild-type $^A\gamma$pro (−299 to +36) in *Hin*cII-cut pUC9. *Nco*I (which cuts $^A\gamma$pro at −141) and *Bam*HI (which cleaves in the polylinker) were used to generate a 166-bp $^A\gamma$pro (−141 to −299) fragment, which was gel purified. To obtain this fragment as a probe, the plasmid was first cut with *Bam*HI, labeled with Klenow fragment, and then recut with *Nco*I. The plasmids pΔ84 and OSS 55–58 were a gift from R. Kingston. Both plasmids (pΔH series) contain the human HSP70 promoter from −84 to *Bam*HI$_{+150}$; however, OSS 55–58 has an oligo scan mutation from −55 to −58. The 174-bp *Hin*dIII$_{-84}$-*Pst*I$_{+90}$ fragment was gel purified and labeled at the *Hin*dIII end with Klenow fragment for use as a probe. Plasmid pFrCK, a gift from E. Golemis and N. Hopkins, contains the Friend murine leukemia virus long terminal repeat from the *Cla*I$_{7676}$-*Kpn*I$_{8301}$ fragment in pCDV1 (numbering starts with the 5′ end of the viral genome). The 175-bp *Cla*I$_{7676}$-*Pst*I$_{7851}$ fragment was gel purified as above. For use as a probe, the fragment was labeled at the *Cla*I end with Klenow fragment.

Cell lines and extracts. Exponentially growing MEL cells were induced with 1.5% dimethyl sulfoxide (DMSO) in Dulbecco modified Eagle medium supplemented with 15% fetal calf serum. Nuclear extracts were made by the procedure of Dignam et al. (14) from the following cell types: uninduced MEL (745-PC4; superinducible subclone B_1-2A); 24-h DMSO-induced MEL; 48-h DMSO-induced MEL; CB5 (a mouse BFU-E [burst-forming unit for erythropoiesis]-like cell line from T. Mak); K562 (a human chronic myelogenous leukemia cell line from D. Tuan and I. London); KMOE-2/05 (a human adult erythremia cell line from A. Menon and J. Lingrel); HL60 (a human promyelocytic cell line from H. von Melchner); and HeLa (a human epithelial carcinoma line from P. Sharp). Protein concentration determinations were done by using the Bio-Rad assay with the bovine gamma globulin standard from Bio-Rad Laboratories.

The following nuclear and whole-cell extracts were generously given by the individuals indicated: 3T3-F442A nuclear extract (mouse adipocyte) from R. Distel and B. Spiegelman; L-cell nuclear extract, human retinoblastoma nuclear extract, 414G (TK^- NIH 3T3) nuclear extract (mouse fibroblast), 70Z/3 nuclear extract (mouse pre-B cell), and PD nuclear extract (mouse pre-B cell) from L. Staudt and D. Baltimore; HAFTL nuclear extract (very early pre-B cell) from D. Weaver and D. Baltimore; F9 whole-cell extract (undifferentiated mouse embryonal carcinoma cells) and BALB/c 3T3 whole-cell extract (mouse embryo fibroblast) from A. Baldwin and P. Sharp; WEHI-231 nuclear extract (mouse B cell) from H. Singh and P. Sharp; L691 nuclear extract (mouse T cell) and BW nuclear extract (mouse T cell) from D. Mbangkallo and N. Hopkins; EL4 nuclear extract (mouse T-cell lymphoma) from N. Manley and N. Hopkins; rat L_6 myoblast nuclear extract and rat L_6 myotube nuclear extract from K. Walsh and P. Schimmel; and WEHI-3 nuclear extract (mouse myelomonocyte) and γ-interferon-induced WEHI-3 nuclear extract from M. Blanar.

Heparin-Sepharose chromatography. Nuclear extract was made as described above from 9.4×10^9 uninduced MEL cells. The yield was 15.3 ml of nuclear extract at 20 mg of protein per ml. A portion of this extract (8 ml) was diluted with 12 ml of buffer A (20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid]-NaOH [pH 7.9], 20% glycerol, 1 mM EDTA, 1 mM dithiothrietol, 1 mM phenylmethylsulfonyl fluoride) to 0.04 M KCl and 8 μg of protein per ml and loaded onto an 8-ml heparin-Sepharose column (Pharmacia) preequilibrated with buffer A plus 0.04 M KCl at a flow rate of 1 column volume per h. The column was

washed with an additional 3 column volumes of buffer A plus 0.04 M KCl and 5 column volumes each of buffer A plus 0.2 M KCl, buffer A plus 0.4 M KCl, and buffer A plus 1 M KCl. The column fractions from each salt wash containing the protein peak were pooled to generate four fraction pools: heparin-Sepharose fraction A (HS-FxnA), 0.04 M KCl; fraction B, 0.2 M KCl; fraction C, 0.4 M KCl; and fraction D, 1 M KCl. The fraction pools were dialyzed against buffer D (20 mM HEPES-NaOH [pH 7.9], 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 1 mM dithiothreitol, 1 mM phenylmethysulfonyl fluoride), and aliquots were stored in liquid N_2.

Gel electrophoresis DNA-binding assay. Binding reactions were carried out in 15 μl containing 5,000 to 20,000 cpm of an end-labeled double-stranded DNA fragment, 1 μl of 0 to 2 μg of poly(dI-dC) · (dI-dC) per μl (Pharmacia) in TE (10 mM Tris [pH 8.0], and 1 mM EDTA) plus 0.1 M NaCl, binding buffer (10 mM Tris hydrochloride [pH 7.5], 1 mM EDTA, 1 mM dithiothreitol, 5% glycerol), 2 to 5 μl of protein extract in buffer D, and additional NaCl or $MgCl_2$ or both as indicated in the figure legends. After the addition of extract, the reaction mixtures were incubated at room temperature for 15 min before the addition of 1 μl of 0.1% xylene cyanol–0.1% bromphenol blue–5% glycerol and a further 10-min incubation (total time, 25 min). The samples were loaded onto low-ionic-strength 4% polyacrylamide gels (31 [width] by 25 [length] by 0.15 cm; acrylamide-bisacrylamide ratio of 29:1). Running buffer (6.7 mM Tris acetate [pH 7.5], 3.3 mM sodium acetate, 1 mM EDTA) was circulated between compartments. Gels were preelectrophoresed for 2 h at 400 V and then for 2 h at 250 V. Gels were electrophoresed at 250 V (10 V/cm) until bromphenol blue had run approximately 15 cm (~2.7 h). The gels were then transferred to Bio-Rad thick blotting filter paper, dried, and autoradiographed.

DNase I footprint. End-labeled DNAs were incubated in preparative binding reaction mixtures that were scaled up six- to sevenfold and that included either 143.5 μg of uninduced MEL HS-FxnA, 1.4 μg of poly(dI-dC) · (dI-dC), 3 mM $MgCl_2$, and 40 mM Na^+ plus K^+ or 124.8 to 145.6 μg of HL60 nuclear extract, 12 to 14 μg of poly(dI-dC) · (dI-dC), 3 mM $MgCl_2$, and 20 mM Na^+ plus K^+ in binding buffer. After 15 min at room temperature, the samples were treated with DNase I (Worthington Diagnostics) at 20 μg/ml (HL60 nuclear extract) or 60 μg/ml (uninduced MEL HS-FxnA) for 15 s at room temperature. Reactions were stopped by the addition of EGTA [ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid] (pH 8.0) to 1 mM to inhibit the DNase I, and the samples were immediately loaded and electrophoresed as described above. Free and complexed DNAs were visualized by autoradiography of the wet gel at 4°C for 4 h with a Du Pont intensifying screen and then were excised from the gel and embedded in a 1% agarose gel containing 0.1% sodium dodecyl sulfate. The labeled DNAs were electrophoresed onto a strip of DEAE membrane (Schleicher & Schuell, Inc.) and eluted at 65°C in TE containing 1 M NaCl. The eluted DNAs were ethanol precipitated twice, rinsed with 95% ethanol, and dried. Samples were suspended in 80% formamide–SeqTBE (134 mM Tris [pH 8.7], 44 mM boric acid, 0.25 mM EDTA)–0.05% xylene cyanol–0.025% bromphenol blue and heated at 90°C for 5 min before analysis by electrophoresis on an 8% acrylamide–7 M urea gel in SeqTBE buffer. The untreated end-labeled DNAs were also subjected to Maxam and Gilbert (31) sequence reactions for G, A+G, and C+T to serve as markers on the sequencing gel. Gels were autoradiographed with preflashed Kodak XAR-5 film at room temperature.

Methylation interference. End-labeled DNA fragments were partially methylated at the N7 of guanine and the N3 of adenine residues as detailed by Maxam and Gilbert (31) with the following modification. The reaction was quenched with 1.5 M sodium acetate (pH 7.0), 1 M β-mercaptoethanol, and 200 μg of glycogen per ml. Methylated DNA was ethanol precipitated twice, rinsed with 95% ethanol, dried, and suspended in TE. The preparative binding reactions were scaled up 13-fold and included 266.5 μg of uninduced MEL HS-FxnA, 2.6 μg of poly(dI-dC) · (dI-dC), 3 mM $MgCl_2$, and 40 mM Na^+ plus K^+ in binding buffer. Binding and gel electrophoresis were performed as described above. Visualization, elution, purification, and denaturing-gel analysis of complexed and free DNAs were performed as described for DNase I footprint analysis, except that the eluted DNAs were cleaved with piperidine (31) and ethanol precipitated twice before the products were resolved by denaturing-gel electrophoresis.

RESULTS

Detection of factors that bind to sequences from β-globin IVS2. We wished to determine whether DNA sequences in the vicinity of the DNase I-hypersensitive site in the IVS2 of β-globin exhibit specific in vitro binding to nuclear proteins isolated from uninduced MEL cells. To screen for such binding, end-labeled DNA fragments were incubated with nuclear extract of MEL cells in the presence of an excess of the synthetic copolymer poly(dI-dC) · (dI-dC) to decrease non-sequence-specific binding of proteins. DNA-protein complexes were resolved from free DNA by electrophoresis in a low-ionic-strength, nondenaturing polyacrylamide gel.

End-labeled subfragments Mβg224-A (+465 to +602) and Mβg224-B (+603 to +689) (Fig. 1) were incubated with nuclear extract made from uninduced MEL cells in the absence or presence of four different unlabeled competitor DNAs to test the sequence specificity of any observed DNA-protein complexes (Fig. 2). Mβg224-A formed several DNA-protein complexes when incubated with MEL cell extract (lane 1). We focused on the two major complexes, which we have designated B1 and B2. Both of these complexes exhibited sequence specificity based on competition experiments (lanes 1 to 5). Mβg224-B formed a single major complex, designated B (lane 9). Complex B also exhibited sequence specificity in competition experiments (lanes 9 to 13). For each of the complexes, the intensity of the band containing the sequence-specific DNA-protein complex was selectively decreased by the inclusion of specific unlabeled competitor DNA Mβg224 in the reaction mixture but not by the inclusion of the nonspecific unlabeled competitor DNA Mβg134, Mβg391, or pUC322 (see Fig. 1 for a description of the Mβg fragments; pUC322 is a 322-bp *Pvu*II fragment from pUC19).

Complex B1 uniformly exhibited a sequence-specific response in the competition assay. Complex B2, however, exhibited variability in response in the competition assay (regardless of the buffer salt concentrations); the complex was sometimes as sensitive to the addition of unlabeled nonspecific DNA as to the addition of unlabeled specific DNA. Since complex B2 is sequence specific (see Fig. 6 to 8), we suggest that the variability of the response of complex B2 with nonspecific competitors may relate to the relatively low salt optima for the formation of this complex (see Fig. 4) and the addition of small amounts of salt with competitor DNAs.

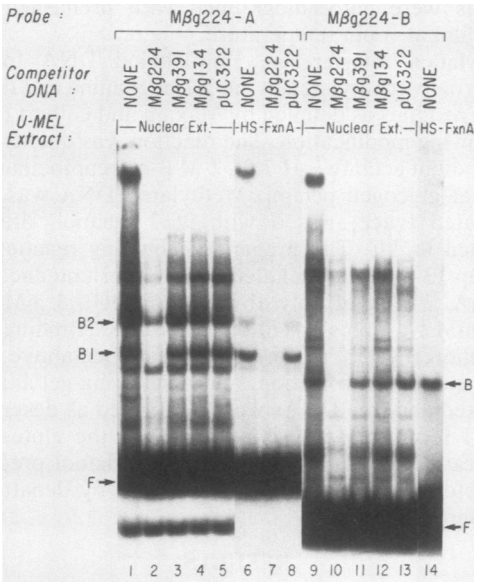

FIG. 2. Detection of nuclear proteins binding to two murine β^M-globin IVS2 DNA segments in uninduced MEL (U-MEL) nuclear extract. End-labeled 156-bp Mβg224-A (~0.45 ng) and 89-bp Mβg224-B (~0.46 ng) fragments were each incubated with 10 μg of uninduced MEL nuclear extract or 8.2 μg of uninduced MEL HS-FxnA (40 mM KCl wash). Each reaction mixture also contained 0.5 μg of poly(dI-dC) · (dI-dC), 20 mM Na^+ plus K^+, 5 mM $MgCl_2$, and various unlabeled DNA fragments in binding buffer. The reaction mixtures were electrophoresed through a low-ionic-strength 4% polyacrylamide gel as described in Materials and Methods. The binding reaction mixtures included: no unlabeled DNA competitor (lanes 1, 6, 9, and 14); 40 ng of unlabeled DNA fragment Mβg224, Mβg391, Mβg134, or pUC322 (lanes 2 to 5 and 10 to 13, respectively); and 20 ng of unlabeled DNA fragment Mβg224 or pUC322 (lanes 7 and 8, respectively). The murine β^M-globin (Mβg) DNA fragments are depicted in Fig. 1, and pUC322 is a 322-bp *Pvu*II fragment from pUC19. F, Migration positions of the free fragments. B, B1, and B2, Migration positions of the bound complexes.

Uninduced MEL cell nuclear extract was fractionated on a heparin-Sepharose column by step elution. The four fractions obtained were tested in binding experiments similar to those described above. HS-FxnA (0.04 M KCl wash) was observed to contain the protein(s) involved in the formation of both the DNA-protein complexes B1 and B2 with Mβg224-A (Fig. 2, lanes 6 to 8) and also of complex B with Mβg224-B (lane 14).

Initial localization of binding. To delimit the binding region within Mβg224-A, oligo2 was synthesized (Fig. 3C). This oligonucleotide was an effective competitor for the protein which forms complex B1 (Fig. 3A, lanes 1 to 4). Owing to the difficulties stated above in interpreting competition for the B2 complex, we cannot say that oligo2 is a specific competitor for the B2 complex formed with Mβg224-A. However, oligo2 formed both complexes B1 and B2 when incubated with the MEL cell nuclear extract HS-FxnA (lanes 9 to 12). These results indicate that this oligonucleotide contains the complete binding site necessary to form both complexes. Oligo2 was also an effective competitor for the formation of complex B by Mβg224-B (lanes 5 to 7). Within oligo2 there is a 17-bp region with 76% homology between DNA fragments Mβg224-A and Mβg224-B (Fig. 3B), suggesting the possibility that this sequence is involved in the formation of either complex B1 or B2 and complex B. The uniform behavior of complex B in the competition assay suggests that it is more likely to correspond to complex B1 than to complex B2. In addition, unlabeled Mβg224-B is an effective specific competitor for the formation of complex B1 on labeled oligo2 (data not shown).

Characterization of salt optima for complexes B1 and B2. To more precisely define the optimal salt concentrations for

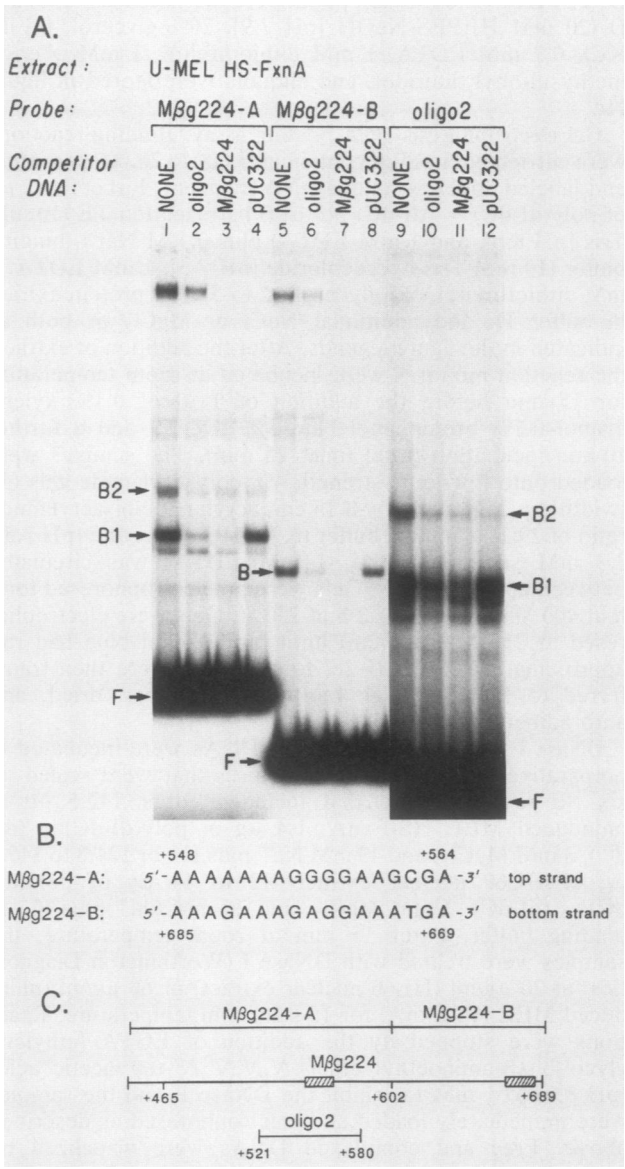

FIG. 3. Region of homology between Mβg224-A and Mβg224-B contains a binding site in common. (A) End-labeled Mβg224-A (0.49 ng), Mβg224-B (0.41 ng), and oligo2 (0.16 ng) were each incubated with 12.3 μg of uninduced MEL (U-MEL) HS-FxnA and 0.2 μg of poly(dI-dC) · (dI-dC) in binding buffer containing 27 mM Na^+ plus K^+ and 5 mM $MgCl_2$. For each labeled DNA fragment, the reaction mixture also contained: no unlabeled DNA competitor (lanes 1, 5, and 9) or unlabeled DNA fragment oligo2 (lanes 2, 6, and 10), Mβg224 (lanes 3, 7, and 11), or pUC322 (lanes 4, 8, and 12). The molar excess of unlabeled DNA fragment ranged from 21- to 30-fold but was consistent for a particular labeled DNA fragment. F, Migration positions of the free fragments. B, B1, and B2, Migration positions of the bound complexes. (B) The 17-bp region of homology between Mβg224-A and Mβg224-B. (C) Diagram of the globin DNAs used as both probes and competitors. ▨, Homologous regions. The base pair numbers are relative to the cap site at +1.

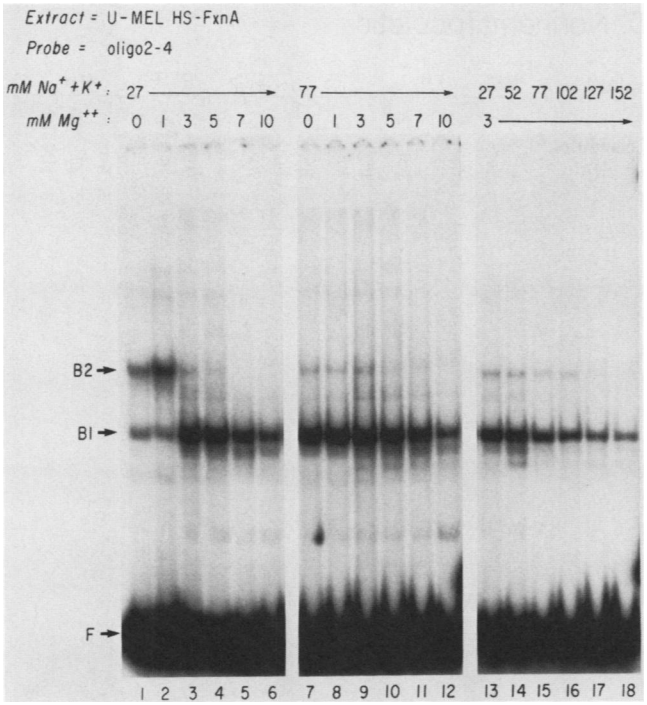

FIG. 4. Optimization of salt conditions for protein binding. Subclone oligo2-4 (see Fig. 8) was excised from pUC19 with *Ava*I-*Eco*RI and end labeled by using Klenow fragment and [α-^{32}P]dATP. Probe oligo2-4 (0.15 ng) was incubated with 12.3 μg of uninduced MEL (U-MEL) HS-FxnA and 0.2 μg of poly(dI-dC) · (dI-dC) in reaction mixtures with various concentrations of Na^+ plus K^+ and $MgCl_2$: 27 mM Na^+ plus K^+ and increasing concentrations of $MgCl_2$ as indicated (lanes 1 to 6); 77 mM Na^+ plus K^+ and increasing concentrations of $MgCl_2$ as indicated (lanes 7 to 12); and increasing concentrations of Na^+ plus K^+ as indicated and 3 mM $MgCl_2$ (lanes 13 to 18). F, Migration position of the free fragment. B1 and B2, Migration positions of the bound complexes.

the formation of complexes B1 and B2, their formation was examined at various concentrations of total Na^+ plus K^+ and $MgCl_2$ (Fig. 4). The optimal salt conditions for the two complexes were very different. Complex B1 formed most efficiently at low total Na^+ plus K^+ and 3 mM $MgCl_2$ (lanes 1 to 6 and 13 to 18), but its stability could be enhanced at $MgCl_2$ concentrations of <3 mM if the total Na^+ plus K^+ was raised above 27 mM (lanes 7 and 8). Complex B2, on the other hand, was most stable at low total Na^+ plus K^+ and 0 to 1 mM $MgCl_2$ (lanes 1 to 6). These differences in salt optima, along with the different electrophoretic mobilities of complexes B1 and B2, suggest that two different proteins are involved in the formation of these two DNA-protein complexes.

Tissue specificity of factors B1 and B2. If either complex B1 or B2 is relevant to tissue specificity of expression of the β-globin gene, then we might expect to observe a limited tissue distribution of the nuclear proteins which form these complexes. To examine this question, nuclear extracts prepared from a variety of cell lines were screened for the ability to form DNA-protein complexes B1 and B2 (Fig. 5 and Table 1). The erythroid cell lines tested included lines with the capacity to express globin genes under the proper stimuli (Fig. 5A). To determine if these factors exhibit major differences in activity during MEL cell differentiation, extracts made from MEL cells treated with DMSO, an inducer of erythroid differentiation, were compared with extracts prepared from uninduced cells (lane 1). Both DNA-protein complexes were observed at significant levels in extracts prepared from MEL cells exposed to DMSO for 24 h (lane 2) and 48 h (lane 3). While precise quantitative comparison is difficult, it is clear that major differences in the levels of formation of complexes B1 and B2 were not observed among these extracts.

Two human erythroid lines, K562 (lane 7) and KMOE (lanes 8 to 10), exhibited formation of complexes B1 and B2 (Fig. 5A). With these extracts, complex B2 appeared as a weak band below a stronger nonspecific band. However, when these extracts were incubated in binding reaction mixtures with optimal salt concentrations for the formation of complex B2, complex B2 was much more evident (lanes 12 and 13). CB5, a mouse BFU-E-like line, appeared to contain factor B2 but not factor B1 (lanes 4 to 6).

Nuclear extracts from a number of nonerythroid hematopoietic cell lines were also tested for the presence of the protein factors responsible for the formation of complexes B1 and B2 (Fig. 5B). None of these extracts contained the protein factor responsible for the formation of the DNA-protein complex B2. This was further confirmed by similar experiments performed at the optimal salt condition for complex B2 formation (data not shown). Still no B2 complex formation was observed with any of these extracts. However, the DNA-protein complex B1 was detected in extracts made from cell lines of B-cell (lanes 2 to 7) and myelomonocyte (lanes 8 to 12) lineages. Very little B1 complex was evident with the HL60 extract (lanes 10 to 12). Control experiments have suggested that significant proteolysis may have occurred during preparation of this extract (data not shown). However, the myelomonocyte lines (HL60 and WEHI-3) also contained a sequence-specific DNA-protein complex (designated B1*) with much faster mobility than complex B1. As discussed below, complex B1* involved the same DNA-binding site as complex B1 and therefore is likely to be related in structure to B1. A sequence-specific complex migrating slightly more slowly than complex B1* was weakly evident with the B-cell extracts 70Z/3 and HAFTL. A similar complex was also occasionally observed with erythroid extracts (Fig. 5A). This complex may result from binding of a proteolytic fragment of factor B1; however, we have not studied these weak complexes further. Extracts made from cell lines of T-cell lineage (lanes 13 to 15 and Table 1) were not observed to form complex B1.

Extracts from several nonhematopoietic cell lines were also examined in the binding assay (Fig. 5C). Neither DNA-protein complex was detected in any of the extracts from these cell lines: F9 embryonal carcinoma (lanes 2 to 4), BALB/c 3T3 fibroblast (lanes 5 to 7), 3T3-F442A adipocyte (lanes 8 to 10), human retinoblastoma (lanes 11 to 13), rat L_6 myotube (lanes 14 to 16), and rat L_6 myoblast, HeLa epithelial carcinoma, L-cell fibroblast, and 414G 3T3 fibroblast (Table 1). A weak complex was observed with the F9 cell extract that comigrated with complex B1, but it was not specific in the competition assay. In control experiments with other probes, it has been shown that all of the extracts tested in this survey (with the possible exception of HL60) contain other expected DNA-binding proteins (e.g., NF-A1, NF-μE1, NF-μE3, CTF, aP2, MLTF, MAPF, and E2F) in intact form and are, therefore, not degraded (2, 7, 15, 40, 46, 49a, 50; D. Galson, unpublished data).

The results of the tissue survey for the factors involved in the formation of complexes B1 and B2 are summarized in Table 1. The DNA-protein complex B2 was only detected in extracts from erythroid-lineage cell lines. The DNA-protein

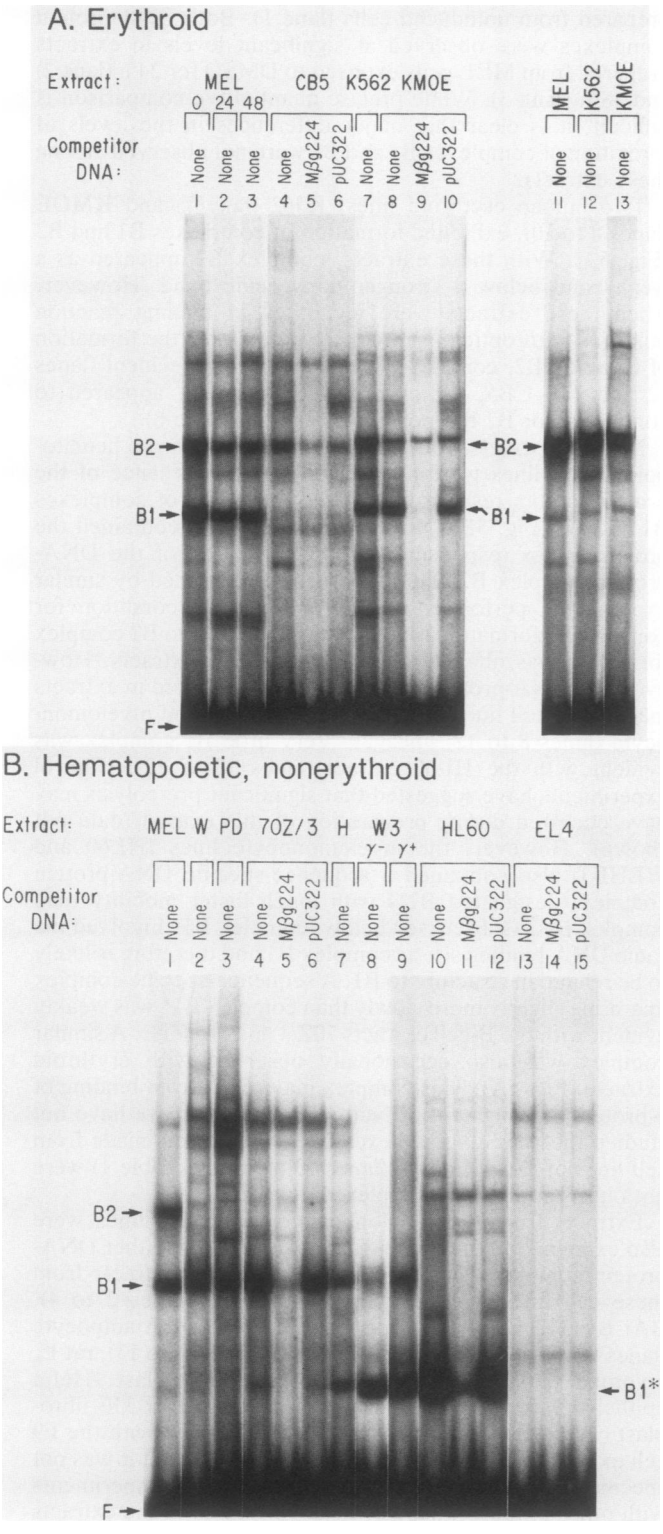

FIG. 5. Analysis of B1 and B2 binding in extracts from a variety of cell lines. (A) Erythroid lines. End-labeled oligo2-4 (0.15 to 0.22 ng) was incubated with various nuclear extracts (9.7 to 17.5 μg) in binding buffer containing 1 to 2 μg of poly(dI-dC) · (dI-dC), 20 mM Na^+ plus K^+, and 3 to 5 mM $MgCl_2$ (lanes 1 to 10) or 1 mM $MgCl_2$ (lanes 11 to 13). Nuclear extract in lanes: 1, uninduced MEL; 2, 24-h DMSO-induced MEL (26% committed cells); 3, 48-h DMSO-induced MEL (63% committed cells); 4 to 6, CB5 (mouse BFU-E-like cell line); 7, K562 (human chronic myelogenous leukemia line); 8 to 10, KMOE (human adult erythemia line); 11, uninduced MEL; 12, K562; 13, KMOE. In some cases, 20× molar excess unlabeled competitor DNAs were also added: Mβg224 (lanes 5 and 9) and pUC322 (lanes 6 and 10). (B) Hematopoietic, nonerythroid lines. End-labeled oligo2-4 (0.16 ng) was incubated with various nuclear extracts (3 to 10 μg) in binding buffer containing 0.5 to 2 μg of poly(dI-dC) · (dI-dC), 70 mM Na^+ plus K^+, and no $MgCl_2$. Nuclear extract in lanes: 1, uninduced MEL; 2, WEHI-231 (W; B-cell); 3, PD (pre-B cell); 4 to 6, 70Z/3 (pre-B cell); 7, HAFTL (H; very early pre-B cell); 8, uninduced WEHI-3 (W3 γ^-; myelomonocyte); 9, γ-interferon-induced WEHI-3 (W3 γ^+); 10 to 12, HL60 (human promyelocytic cell); 13 to 15, EL4 (T-cell lymphoma). In some cases, 20X molar excess unlabeled competitor DNAs were also added: Mβg224 (lanes 5, 11, and 14) and pUC322 (lanes 6, 12, and 15). All the cell lines were derived from mouse cells except HL60. (C) Nonhematopoietic lines. End-labeled oligo2-4 (0.15 to 0.22 ng) was incubated with various nuclear extracts (3 to 10 μg) in binding buffer containing 0.2 to 2 μg of poly(dI-dC) · (dI-dC) and either 20 mM Na^+ plus K^+ plus 3 to 5 mM $MgCl_2$ or 70 mM Na^+ plus K^+ plus no $MgCl_2$. Extract in lanes: 1, uninduced MEL nuclear extract; 2 to 4, F9 whole-cell extract (mouse embryonal carcinoma cell); 5 to 7, BALB/c 3T3 whole-cell extract (mouse fibroblast); 8 to 10, 3T3-F442A nuclear extract (Adip; mouse adipocyte); 11 to 13, human retinoblastoma nuclear extract (RetBl); 14 to 16, L_6 nuclear extract (Myo; rat myotube). In some cases, 20X molar excess unlabeled competitor DNAs were also added: Mβg224 (lanes 3, 6, 9, 12, and 15) and pUC322 (lanes 4, 7, 10, 13, and 16). All of the extracts tested in panels A, B, and C and in Table 1 have been shown to contain other characterized DNA-binding proteins in intact form and are therefore not degraded (data not shown). F, Migration positions of the free fragments. B1*, B1, and B2, Migration positions of the bound complexes.

complex B1 was detected in extracts from erythroid-lineage cell lines, B-cell-lineage cell lines, and myelomonocyte-lineage cell lines but not T-cell-lineage cell lines or any nonhematopoietic cell lines.

Binding domains for B1 and B2. DNase I footprint analysis (19) was used to delineate the binding domains of the factors

TABLE 1. Tissue specificity of binding factors

Cell line[a]	Cell type	Complex B1	Complex B2
Erythroid			
Uninduced-MEL	Murine proerythroblast	+	+
24I-MEL	Murine normoblast	+	+
48I-MEL	Murine normoblast	+	+
CB5	Murine BFU-E-like	−	+
K562	Human chronic myelogenous leukemia	+	+
KMOE	Human adult erythremia	+	+
Hematopoietic			
WEHI-231	Murine B cell	+	−
PD	Murine pre-B cell	+	−
70Z/3	Murine pre-B cell	+	−
HAFTL	Murine very pre-B cell	+	−
WEHI-3	Murine myelomonocyte	+	−
HL60	Human promyelocyte	+	−
EL4	Murine T-cell lymphoma	−	−
L691*	Murine T-cell lymphoma	−	−
BW*	Murine T-cell lymphoma	−	−
Nonhematopoietic			
HeLa*	Human cervical carcinoma	−	−
F9	Murine embryonal carcinoma	−	−
BALB/c 3T3	Murine fibroblast	−	−
L-cell*	Murine fibroblast	−	−
414G 3T3*	Murine fibroblast	−	−
RetBl	Human retinoblastoma	−	−
L_6 MYOT	Rat myotube	−	−
L_6 MYOB*	Rat myoblast	−	−
3T3-F442A	Murine adipocyte	−	−

[a] The data for this table were obtained from Fig. 5 except for cell lines designated with an asterisk (*) (data not shown). 24I-MEL and 48I-MEL, Cells induced with DMSO for 24 and 48 h, respectively.

present in complexes B1 and B2 (Fig. 6). Binding reactions containing either uninduced MEL nuclear extract HS-FxnA or HL60 nuclear extract were treated with DNase I to produce a partial digestion of the DNA. The B1 and B2 complexes and free DNA were resolved by electrophoresis on a low-ionic-strength native polyacrylamide gel. The DNA was eluted from the bands containing complexes B1 and B2 and free DNA and analyzed on a sequencing gel. DNAs from G, A+G, and C+T Maxam and Gilbert sequencing reactions were coelectrophoresed to provide identification of the region protected from DNase I cleavage. The region of DNase I protection was different for complexes B1 and B2. However, the DNase I-protected region of complex B1 formed with extract derived from uninduced MEL cells was the same as that of complex B1* observed with extract derived from HL60 cells. In complex B1, the protected region encompassed 17 bp (+551 to +567) on the coding strand and 16 bp (+549 to +564) on the noncoding strand in extracts derived from both uninduced MEL (Fig. 6A and B, lanes 2) and HL60 (lanes 4) cells. This region overlapped the region of homology with DNA fragment Mβg224-B (Fig. 3B), confirming the suggestion from the data in Fig. 3 that it is complex B1 observed with Mβg224-A that corresponds to complex B observed with Mβg224-B. The two B1-binding sites in Mβg224-A and Mβg224-B were therefore designated B1-A and B1-B, respectively. In complex B2 (observed only in erythroid lines), the protected region extended over 8 bp (+539 to +546) on the coding strand (Fig. 6A, lane 1) and 12 bp (+537 to +548) on the noncoding strand (Fig. 6B, lane 1). The binding domains for complexes B1 (B1-A site) and B2 are next to each other, which suggests the possibility that the proteins involved in the formation of the two complexes interact with each other. The extent of protection of nucleotides within the protected regions was not uniform, and in fact, there were some nucleotides within the protected region of complex B1 (+559 coding; +553 and +556 noncoding) which exhibited enhanced sensitivity to DNase I cleavage when protein was bound. This suggests that the interaction of the protein with the DNA results in an alteration in the structure of the DNA helix.

Sequence of the binding sites for B1 and B2. A methylation interference assay was used (44) to define more precisely the contact residues on the DNA for the two binding factors. This assay identifies methylations both at the N7 of guanine residues (in the major groove) and at the N3 of adenine residues (in the minor groove) which prevent binding of a factor to the DNA. DNA molecules with a *methyl*-G or *methyl*-A that interferes with binding are depleted from the population in the DNA-protein complex, and this results in gaps in the sequence ladder of the complexed DNA compared with that of free DNA. For this assay, probe oligo2-4, end labeled on either one strand or the other, was partially methylated by dimethyl sulfate and incubated with uninduced MEL nuclear extract HS-FxnA. The resultant DNA in complexes B1 and B2, as well as the free DNA, was isolated from a preparative gel, cleaved with piperidine, and resolved on a sequencing gel (Fig. 7). In complex B1, numerous contacts in both the major and minor grooves are evident on the coding strand (Fig. 7A, lane 5) and are all contained within the DNase I-protected region mapped in Fig. 6. However, owing to the presence of mostly cytidine and thymidine residues on the noncoding strand, limited information can be deduced, except to exclude interference by methylation of the one guanine (at +562) within the DNase I-protected region on that strand (Fig. 7B, lane 5). In complex B2, the interfering adenines and guanines are again

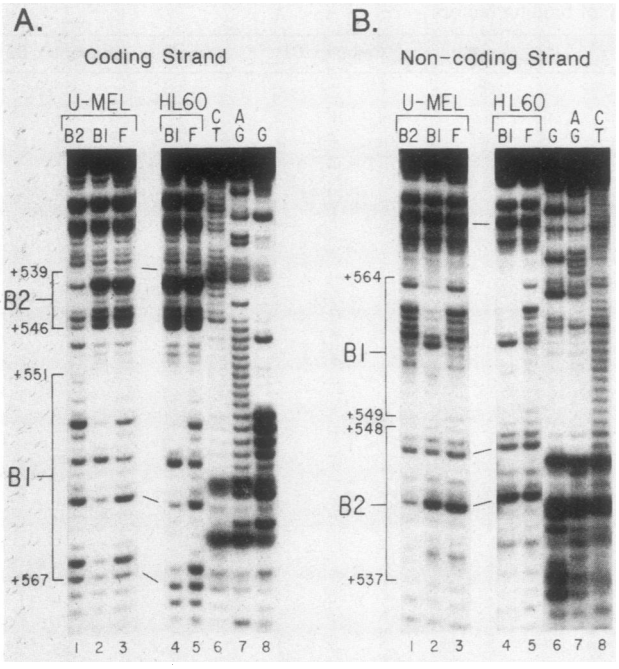

FIG. 6. DNase I footprint analysis of the B1 and B2 complexes. DNase I footprinting was performed as described in Materials and Methods. *Ava*I-*Eco*RI-excised oligo2-4 was labeled by Klenow fragment on the coding strand (A) at the *Eco*RI end with [α-^{32}P]dATP and on the noncoding strand (B) at the *Ava*I end with [α-^{32}P]dCTP. DNase I protection assays were carried out on the labeled DNAs in binding reactions containing either uninduced MEL (U-MEL) HS-FxnA (lanes 1 to 3) or HL60 nuclear extract (lanes 4 and 5), and the DNA-protein complexes were separated from free DNA (F) by gel electrophoresis as described in Materials and Methods. In the case of reactions containing uninduced MEL HS-FxnA, DNA from the complexes designated B2 (lanes 1) and B1 (lanes 2) and free DNA (lanes 3) were eluted. In the case of reactions containing HL60 nuclear extract, DNA from the complex designated B1* (lanes 4) in Fig. 5 and free DNA (lanes 5) were eluted. The eluted DNAs were analyzed by electrophoresis on sequencing gels. Maxam and Gilbert sequence reactions for C+T, A+G, and G were used as size markers (lanes 6 to 8, as indicated above each lane). The brackets summarize the DNase I protection. The base pair numbers are relative to the cap site at +1.

all contained within the DNase I-protected region mapped in Fig. 6. Multiple contacts were observed on the noncoding strand (Fig. 6B, lane 4) and with the adenines (at +543 and +546) on the coding strand (Fig. 6A, lane 4).

A summary of the data obtained from both the DNase I protection and methylation interference studies is presented in Fig. 7C. In general, interference was weaker with the guanine and adenine residues at the borders of the binding sites. The binding sites as defined by methylation interference were several bases smaller on either side than were the DNase I-protected regions. The proposed binding sites for the two complexes are 5′-AAAGGGGAAGCG-3′ for B1-A and 5′-TCCTATCA-3′ for B2.

Sequence specificity of factors B1 and B2. To characterize more precisely the base sequence specificity of the binding interactions of B1 and B2, four isolates of subcloned oligo2 that contained sequence mutations (Fig. 8A) were tested in the electrophoresis mobility shift assay (Fig. 8B). This provided further insight into the binding-site requirements for complexes B1 and B2. The probe in lane 1 is the oligo2 mixture prior to subcloning; both complexes B1 and B2 are evident. The probes in lanes 2 to 5 are the subclones of oligo2: 2-1, 2-4, 2-10, and 2-14, respectively. All of the subclones could form complex B1, even though oligo2-14 (lane 5) contains a G-to-A change at +555, a site within the B1-binding domain. However, the single-base deletions within the binding site for complex B2 in oligo2-1 and oligo2-10 negatively affected the formation of the B2 complex. Oligo2-1 (lane 2) could participate weakly in complex B2 formation; however, oligo2-10 (lane 4) could not form complex B2 at all. Although it may appear that oligo2-4 (lane

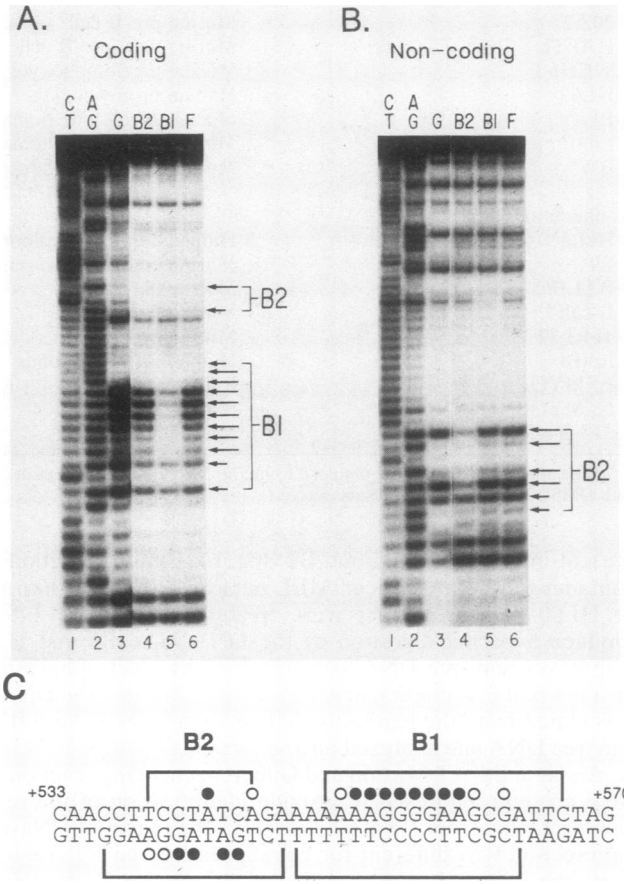

FIG. 7. Methylation interference footprint analysis of the B1 and B2 complexes. Methylation interference footprinting was performed as described in Materials and Methods. *Ava*I-*Eco*RI-excised oligo2-4 was labeled by Klenow fragment on the coding strand (A) at the *Eco*RI end with [α-^{32}P]dATP and on the noncoding strand (B) at the *Ava*I end with [α-^{32}P]dCTP. Both labeled DNAs were partially methylated and incubated in binding reactions containing uninduced MEL HS-FxnA (lanes 4 to 6), and the DNA-protein complexes were separated from free DNA by gel electrophoresis as described in Materials and Methods. DNAs from the complexes designated B2 (lanes 4) and B1 (lanes 5) and free DNA (F; lanes 6) were eluted. The eluted DNAs were cleaved by piperidine, and the products were analyzed by electrophoresis on sequencing gels. Maxam and Gilbert sequence reactions for C+T, A+G, and G were used as size markers (lanes 1, 2, and 3, respectively). The long arrows designate strongly interfering *methyl*-guanine and *methyl*-adenine residues; the short arrows designate weakly interfering *methyl*-guanine and *methyl*-adenine residues. (C) Summary of the DNase I protection results from Fig. 6 and the methylation interference results above. The brackets indicate the DNase I-protected regions. Symbols: ●, strongly interfering *methyl*-guanine and *methyl*-adenine residues; ○, weakly interfering *methyl*-guanine and *methyl*-adenine residues.

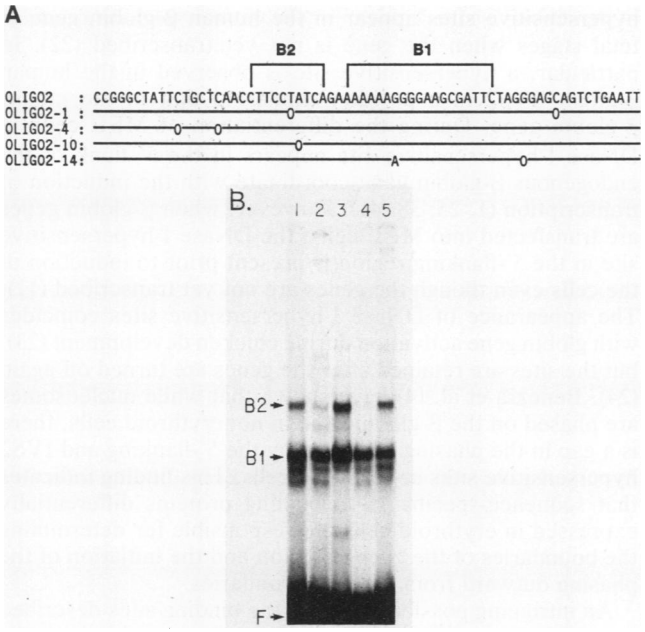

FIG. 8. Binding to four subclones of oligo2 with sequence differences. (A) The sequence of oligo2 and the four subclones, with their nucleotide differences indicated as O to indicate a missing nucleotide or A to indicate a change from the original sequence. The brackets indicate the DNase I-protected regions from both strands combined for each complex as marked. (B) Each of the oligo2 subclones was excised from pUC19 with *Ava*I-*Eco*RI. The subclones and oligo2 were then end labeled by using Klenow fragment and [α-^{32}P]dATP. All probes were purified from a polyacrylamide gel. Each probe was incubated with 12.3 μg of U-MEL HS-FxnA and 0.2 μg of poly(dI-dC) · (dI-dC) in binding buffer containing 27 mM Na^+ plus K^+ and 5 mM $MgCl_2$. Lanes: 1, oligo2; 2, oligo2-1; 3, oligo2-4; 4, oligo2-10; 5, oligo2-14. F, Migration position of the free fragment. B1 and B2, Migration positions of the bound complexes.

3) bound factor B2 more efficiently than did the oligo2 mixture (lane 1), this was not a reproducible observation. Therefore, the mutations in oligo2-4 had no effect on the binding of either B1 or B2.

The mouse β^m-globin IVS2 also contains binding sites B1-A, B1-B, and B2. The B2-binding site is completely homologous between the β^M- and β^m-globin genes. However, there are sequence differences in the B1-binding sites between the β^M- and β^m-globin genes as well as between the B1-A and B1-B sites themselves. The different forms of each of the two B1-binding sites in the β^M- and β^m-globin IVS2 regions, as well as the oligo2-14 mutation, are shown in Fig. 9A. These five sequences have been demonstrated to participate in the formation of complex B1 (data not shown for β^m-globin B1 sites). The differences observed between different B1 sites are all conservative changes: purine for purine and pyrimidine for pyrimidine. A possible core consensus, delimited by the methylation interferences observed in Fig. 7, for the B1-binding site is indicated underneath.

To further examine the binding-site requirements for the formation of the B1 complex, competition assays with DNA fragments containing sequences with homology to the B1-binding site were carried out (Fig. 9B). The unlabeled DNAs were tested for the ability to be an effective competitor for the protein which forms complex B1 on labeled oligo2-4. Also, a 20× molar excess of unlabeled Mβg244 was tested for specific competition against each of the DNAs as probe. The following DNAs with 58 to 75% homology to the 12-bp core consensus sequence for B1 were tested for the ability to bind factor B1: a fragment from the 5′-flanking DNase I-hypersensitive site of mouse β^M-globin; a fragment from the 5′-flanking DNase I-hypersensitive site of human $^A\gamma$-globin (22); the wild-type human HSP70 promoter serum response element (SRE) (52); an oligo scan mutant of the human HSP70 SRE; and the 175-bp *Cla*I$_{7676}$-*Pst*I$_{7851}$ fragment from Friend murine leukemia virus containing the U3 border (the precise labeled and unlabeled DNAs used in this study are described in Materials and Methods). None of the DNAs shown in Fig. 9B could effectively compete for factor B1, nor could they form any B1-like complex. These results indicate that although some nucleotide variation in the B1-binding sites that can form the B1 complex exists, the interaction is quite sequence specific and random purine-rich stretches do not bind factor B1.

DISCUSSION

Using an electrophoretic mobility shift assay, we detected two distinct sequence-specific factors in uninduced MEL nuclear extract that could bind to motifs within the DNase I-hypersensitive region of mouse β-globin IVS2. One of these factors, B1, bound to two homologous sites, B1-A and B1-B, which are approximately 100 bp apart and on opposite DNA strands. The tissue distribution of this factor was limited to a subset of hematopoietic cell lines. The other factor, B2, bound to a site very close to the B1-A site and was limited to cells of the erythroid lineage.

Factor B1 was found in all the erythroid cell types tested except for CB5. This cell line has been determined to be blocked at a very early stage of erythropoiesis and resembles a BFU-E cell (42). The absence of factor B1 in the CB5 cells suggests the possibility that the factor is differentiation-stage specific for later stages. We were unable to demonstrate this point directly, because CB5 cells do not respond to inducers of differentiation in a synchronous manner. As a result, it is not possible to obtain large uniform populations of late-stage erythroid cells from CB5 cultures.

The presence of factor B1 in several nonerythroid hematopoietic cell types suggests that this factor may be important for the control of expression of other genes besides

A. SEQUENCES DEMONSTRATED TO FORM COMPLEX B1:

B1 Site in Oligo2-14		AAAAAAAAGGGAAGCGATTC	
B1-A Site in β-major-globin	+548	AAAAAAAGGGGAAGCGATTC	+567
B1-A Site in β-minor-globin	+548	AAGGAAAGGGGAAGCGATTC	+567
B1-B Site in β-major-globin	+685	AAAGAAAGAGGAAATGACAA	+666
B1-B Site in β-minor-globin	+705	CAAGAAAGAGGAAATGACAA	+686
B1 CORE CONSENSUS		-arrAAARRGGAARYGay--	

B. SEQUENCES WHICH FAIL TO FORM COMPLEX B1:

5' flank mouse β-major-globin	-310	aaaaAAAGGAAAATTA	-293
5' flank human Aγ-globin	-204	ccccCAAGAGGATACT	-219
WT huHSP70 SRE	-61	ctcaGAAGGGAAAAGG	-46
MUT huHSP70 SRE (OSS 55-58)	-62	gctcGACGGGGAAAAG	-47
F-MuLV LTR	7805	aaaaAGGGGGGAATGA	7820

FIG. 9. Sequence specificity of factor B1. (A) The B1-A sites are on the coding strand and the B1-B sites are on the noncoding strand. The methylation interference results from Fig. 7 were used to delimit the B1 core consensus. (B) The underlined nucleotides deviate from the putative B1 core consensus sequence above. Nucleotide positions from the respective cap site (+1) are indicated, except for Friend murine leukemia virus long terminal repeat (F-MuLV LTR), which has the nucleotide position from the 5′ end of the viral genome indicated. See text for details.

β-globin. A very substantial amount of complex B1* was formed by HL60 nuclear extract. The DNase I protection footprint of complex B1* was identical to that of complex B1 (Fig. 6). Therefore, the protein involved in the formation of complex B1* is probably related to factor B1. Because HL60 extracts have high levels of proteolytic activity, we suggest that B1* may derive from B1 by cleavage. Alternatively, factor B1* could represent a different polypeptide from factor B1. The octamer-motif-binding proteins NF-A1 and NF-A2 provide a precedent for different *trans*-acting factors recognizing the same binding site in different cell types (46).

The presence of DNase I protection and methylation interference footprints demonstrates that the formation of complex B2 exhibits strong sequence specificity. This conclusion is further strengthened by the demonstration that the deletion of a single base in the B2-binding site was sufficient to prevent B2 binding (Fig. 8). However, in the competition assays, the ability of unlabeled nonspecific DNAs to compete was sometimes equal to that of the unlabeled specific DNA even at optimal buffer salt concentrations. This observation may be considered in relation to the salt optima for this DNA-protein interaction. This complex requires both very low monovalent cation and very low Mg^{2+} concentrations. Consequently, the complex would be very sensitive to the addition of any positively charged contaminants contained in the unlabeled DNA preparations independent of the DNA sequence content of the competitor DNA.

The erythroid specificity of factor B2 raises some questions. Is this factor involved in the regulation of other erythroid-specific genes? Could factor B2 play a significant role in establishing the restriction of β-globin gene transcription to cells of the erythroid lineage? The latter hypothesis could be tested by site-directed mutagenesis of the B2-binding site to eliminate B2 binding and analysis of the tissue specificity of expression of the mutant β-globin genes.

The footprint data in Fig. 6 and 7 indicate that factors B1 and B2 bound at independent sites. This conclusion is supported by the fact that B1 and B2 binding could be observed separately in certain extracts; i.e., B1 binding, but not B2 binding, was observed in nonerythroid hematopoietic cell extracts, and B2 binding, but not B1 binding, was observed in CB5 extracts. In addition, site B1-B in both the β^M- and β^m-globin genes could be bound by factor B1 even in the absence of an accompanying B2-binding site, and oligo2-1 and oligo2-10, which have mutations in the B2-binding site that negatively affect B2 binding, could still bind factor B1 efficiently.

The B2-binding site delineated by methylation interference footprint analysis extended from +539 to +546 in IVS2, whereas site B1-A included the sequences from +552 to +563. Thus, the two sites are only separated by 5 nucleotides (the DNase I footprints are even closer). The proximity of these two binding sites indicates that the two factors could either interact cooperatively or sterically block one another such that only one can bind at a time. Either possibility could serve a regulatory role. The enhanced DNase I sensitivity of certain bases upon protein binding in complex B1 suggests that the DNA helix structure becomes distorted by the interaction with factor B1. Such bending has been observed with other DNA-protein interactions (28, 32, 37, 43, 49). The stored energy of an induced bend could be used in a regulatory manner, such as opening the DNA helix for transcription.

Hypersensitive sites have been most strongly implicated in the establishment of a permissive conformation that is required, but is not sufficient, for transcription. Several hypersensitive sites appear in the human β-globin gene at fetal stages when the gene is not yet transcribed (22). In particular, a hypersensitive site is observed in the human β-globin IVS2 in a position similar to that in the mouse β-globin gene. During the differentiation of MEL cells, a DNase I-hypersensitive site appears in the 5′ flank of the endogenous β-globin gene coordinate with the induction of transcription (1, 25, 38, 41). However, when β-globin genes are transfected into MEL cells, the DNase I-hypersensitive site in the 5′-flanking region is present prior to induction of the cells even though the genes are not yet transcribed (11). The appearance of DNase I-hypersensitive sites coincides with globin gene activation during chicken development (23), but the sites are retained after the genes are turned off again (24). Benezra et al. (4) have shown that while nucleosomes are phased on the β-globin gene in nonerythroid cells, there is a gap in the phasing bordered by the 5′-flanking and IVS2 hypersensitive sites in erythroid cells. This finding indicates that sequence-specific DNA-binding proteins differentially expressed in erythroid cells are responsible for determining the boundaries of the gapped region and the initiation of the phasing outward from the gap boundaries.

An intriguing possibility is that the binding sites described in this study may play such a role(s). The spacing between the B1-A and B1-B sites of about 100 bp or approximately one turn around a nucleosome is particularly suggestive in this context. This spacing would permit DNA folding in the nucleosome to bring these sites into close proximity in three-dimensional space. Alternatively, B1 factors bound at the two B1 sites could cooperatively interact to cause DNA looping, which has been demonstrated to occur in several systems (36, 39). Although independent binding to each site was detected, the functional roles of the sites may rely on the pairwise relationship. The putative importance of binding site B1-B is supported by the observation that the β^M- and β^m-globin genes are very homologous through to +597 (3′ to B1-A) and then diverge greatly through the rest of IVS2 except for another short span of 43 bp with good homology between them at +659 to +702 β^M and +679 to +722 β^m; this second region contains the B1-B site. In addition, very close to site B1-B (+666 to +685) is a match of seven of eight residues (AGTTGCAT) with the octamer motif (at +704) involved in the enhancer activity of a number of genes (5, 12, 33, 45). We have established that this site can be bound by the ubiquitous factor NF-A1 (D. Galson, unpublished data). There is also a binding site for NF-A1 (an eight-of-eight match) at −322 of the β^M-globin gene (D. Galson, unpublished data), which is in the neighborhood of the 5′ DNase I-hypersensitive site. The octamer motif and the new binding sites identified in this study may all have roles in the creation or maintenance of the DNase I-hypersensitive structure in the IVS2.

Coordinate with the induction of β-globin transcription in MEL cells, the DNase I-hypersensitive site within the IVS2 decreases in intensity, although it does not disappear entirely (1, 4, 25, 38, 41). This change could be the result of loss of the hypersensitive site structure itself or could be due to the presence of transcription complexes covering the hypersensitive site structure and hindering access by DNase I. If factor B1 or factor B2 or both are required in a rate-limiting way for the maintenance of the hypersensitive site structure and the hypersensitive structure is lost when transcription commences, then one would expect the levels of B1 or B2 or both to drop as MEL cells induce. However, if the hypersensitive structure is still present during transcription, then the levels of B1 and B2 should not change as MEL cells

induce. It is clear from the data in Fig. 5A that neither factor was lost by 48 h of DMSO induction. Since we were unable to precisely quantitate the relative levels of B1 and B2 as the cells induced, it is not known whether the levels decrease at all or do so coordinately with the decrease in the intensity of the IVS2 DNase I-hypersensitive site. Consequently, it is still unknown whether the IVS2 DNase I-hypersensitive structure is lost during transcription or precisely what roles the two factors may play.

Although the regulatory role(s) that factors B1 and B2, which bind to intragenic sites, may have in the control of β-globin expression has not been determined, examples exist of intragenic regulatory elements in the immunoglobulin genes (3, 21) and the adenovirus E1A gene (35). The possible roles for factors B1 and B2 include: (i) negative control of β-globin expression in the uninduced MEL cell; (ii) a necessary, but not sufficient, positive role involving establishment of a permissive conformation on the β-globin gene in the uninduced MEL cell; and (iii) a positive requirement for β-globin expression upon the induction of the MEL cell. The identification of binding sites B2, B1-A, and B1-B in the DNase I-hypersensitive region of IVS2 provides a framework for testing the significance of the binding interactions in the control of β-globin gene expression.

ACKNOWLEDGMENTS

We thank Al Baldwin, Richard Carthew, Jim Croop, Jack Hensold, Cynthia Jackson, Rachel Kent, Phillip Sharp, and Harinder Singh for helpful advice and for various reagents. We especially thank Al Baldwin, Jim Croop, and Rachel Kent for critically reading the manuscript. We thank Joan Lane for her careful preparation of DNA fragments. We also thank all who contributed DNAs and extracts.

D.L.G. was supported by Public Health Service grant CA-07533 from the National Cancer Institute. This study was supported by Public Health Service grant CA-17575 to D.E.H. from the National Cancer Institute.

LITERATURE CITED

1. **Balcarek, J. M., and F. A. McMorris.** 1983. DNase I hypersensitive sites of globin genes of uninduced Friend erythroleukemia cells and changes during induction with dimethyl sulfoxide. J. Biol. Chem. **238**:10622–10628.
2. **Baldwin, A. S., Jr., and P. A. Sharp.** 1987. Binding of a nuclear factor to a regulatory sequence in the promoter of the mouse *H-2K*^{b} class I major histocompatibility gene. Mol. Cell. Biol. **7**:305–313.
3. **Banerji, J., L. Olsen, and W. Schaffner.** 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33**:729–740.
4. **Benezra, R., C. R. Cantor, and R. Axel.** 1986. Nucleosomes are phased along the mouse β-major globin gene in erythroid and nonerythroid cells. Cell **44**:697–704.
5. **Bohmann, D., W. Keller, T. Dale, H. R. Scholer, G. Tebb, and I. W. Mattaj.** 1987. A transcription factor which binds to the enhancers of SV40, immunoglobulin heavy chain and U2 snRNA genes. Nature (London) **325**:268–272.
6. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
7. **Carthew, R. W., L. A. Chodosh, and P. A. Sharp.** 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43**:439–448.
8. **Chada, K., J. Magram, and F. Constantini.** 1986. An embryonic pattern of expression of a human fetal globin gene in transgenic mice. Nature (London) **319**:685–689.
9. **Chada, K., J. Magram, K. Rapheal, G. Radice, E. Lacy, and F. Constantini.** 1985. Specific expression of a foreign β-globin gene in erythroid cells of transgenic mice. Nature (London) **314**:377–380.
10. **Charnay, P., P. Mellon, and T. Maniatis.** 1985. Linker scanning mutagenesis of the 5'-flanking region of the mouse β-major-globin gene: sequence requirements for transcription in erythroid and nonerythroid cells. Mol. Cell Biol. **5**:1498–1511.
11. **Charnay, P., R. Triesman, 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.
12. **Davidson, I., C. Fromental, P. Augereau, A. Wildeman, M. Zenke, and P. Chambon.** 1986. Cell-type specific protein binding to the enhancer of simian virus 40 in nuclear extracts. Nature (London) **323**:544–548.
13. **Dierks, P., A. Van Ooyen, M. D. Cochron, C. Dobkin, J. Reiser, and C. Weissman.** 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.
14. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
15. **Distel, R. J., H.-S. Ro, B. S. Rosen, D. L. Groves, and B. M. Spiegelman.** 1987. Nucleoprotein complexes that regulate gene expression in adipocyte differentiation: direct participation of c-*fos*. Cell **49**:835–844.
16. **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.
17. **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.
18. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6506–6525.
19. **Galas, D., and A. Schmitz.** 1978. DNase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. **5**:3157–3170.
20. **Garner, M. M., and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions. Nucleic Acids Res. **9**:3047–3060.
21. **Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa.** 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell **33**:717–728.
22. **Groudine, M., T. Kohwi-Shigematsu, R. Gelinas, G. Stamatoyannopoulos, and T. Papayannopoulou.** 1983. Human fetal to adult hemoglobin switching: changes in chromatin structure of the β-globin gene locus. Proc. Natl. Acad. Sci. USA **80**: 7551–7555.
23. **Groudine, M., and H. Weintraub.** 1981. Activation of globin genes during chicken development. Cell **24**:393–401.
24. **Groudine, M., and H. Weintraub.** 1982. Propagation of globin DNAase I-hypersensitive sites in the absence of factors required for induction: a possible mechanism for determination. Cell **30**:131–139.
25. **Hofer, E., R. Hofer-Warbinek, and J. E. Darnell, Jr.** 1982. Globin RNA termination: a possible termination site and demonstration of transcriptional control correlated with altered chromatin structure. Cell **29**:887–893.
26. **Kollias, G., N. Wrighton, J. Hurst, and F. Grosveld.** 1986. Regulated expression of human ^{A}γ-, β-, and hybrid γβ-globin genes in transgenic mice: manipulation of the developmental expression patterns. Cell **46**:89–94.
27. **LaFlamme, S., S. Acuto, D. Markowitz, L. Vick, W. Landschultz, and A. Bank.** 1987. Expression of chimeric human β- and δ-globin genes during erythroid differentiation. J. Biol. Chem. **262**:4819–4826.
28. **Liu-Johnson, H.-N., M. R. Gartenberg, and D. M. Crothers.** 1986. The DNA binding domain and bending angle of *E. coli* CAP protein. Cell **47**:995–1005.
29. **Magram, J., K. Chada, and F. Constantini.** 1985. Developmental regulation of a cloned adult β-globin gene in transgenic mice.

Nature (London) **315**:338–340.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**: 499–560.
32. **McLarin, J. A., C. A. Frederick, B.-C. Wang, P. Greene, H. W. Boyer, J. Grable, and J. M. Rosenberg.** 1986. Structure of the DNA-Eco R1 endonuclease recognition complex at 3 Å resolution. Science **234**:1525–1541.
33. **Murphy, J. T., J. T. Skuzeski, E. Lund, T. H. Steinberg, R. R. Burgess, and J. E. Dahlberg.** 1987. Functional elements of the U1 RNA promoter. J. Biol. Chem. **262**:1795–1803.
34. **Myers, R., K. Tilly, and T. Maniatis.** 1986. Fine structure genetic analysis of a β-globin promoter. Science **232**:613–618.
35. **Osborne, T. F., D. N. Arvidson, E. S. Tyau, M. Dunsworth-Browne, and A. J. Berk.** 1984. Transcription control region within the protein-coding portion of adenovirus E1A genes. Mol. Cell. Biol. **4**:1293–1305.
36. **Ptashne, M.** 1986. Gene regulation by proteins acting nearby and at a distance. Nature (London) **322**:697–701.
37. **Richmond, T.** 1987. Protein-DNA interaction: another folding problem? Nature (London) **326**:18–19.
38. **Salditt-Georgieff, M., M. Sheffrey, K. Krauter, J. E. Darnell, Jr., R. Rifkind, and P. A. Marks.** 1981. Induced transcription of the mouse β-globin transcription unit in erythroleukemia cells: time course of induction and of changes in chromatin structure. J. Mol. Biol. **172**:437–450.
39. **Schlief, R.** 1987. Why should DNA loop? Nature (London) **327**:369–370.
40. **Sen, R., and D. Baltimore.** 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell **46**: 705–716.
41. **Sheffrey, M., R. A. Rifkind, and P. Marks.** 1982. Murine erythroleukemia cell differentiation: DNAase I hypersensitivity and DNA methylation near the globin genes. Proc. Natl. Acad. Sci. USA **79**:1180–1184.
42. **Shibuya, T., and T. W. Mak.** 1983. Isolation and induction of erythroleukemic cell lines with properties of erythroid progenitor burst-forming cell (BFU-E) and erythroid precursor cell (CFU-E). Proc. Natl. Acad. Sci. USA **80**:3721–3725.
43. **Shuey, D. J., and C. S. Parker.** 1986. Bending of promoter DNA on binding of heat shock transcription factor. Nature (London) **323**:459–461.
44. **Siebenlist, U., and W. Gilbert.** 1980. Contacts between *Escherichia coli* RNA polymerase and early promoter phage T7. Proc. Natl. Acad. Sci. USA **77**:122–126.
45. **Singh, H., R. Sen, D. Baltimore, and P. Sharp.** 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature (London) **319**:154–158.
46. **Staudt, L. M., H. Singh, R. Sen, T. Wirth, P. A. Sharp, and D. Baltimore.** 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature (London) **323**: 640–643.
47. **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.
48. **Townes, T. M., J. B. Lingrel, H. Y. Chen, R. L. Brinster, and R. D. Palmiter.** 1985. Erythroid-specific expression of human β-globin genes in transgenic mice. EMBO J. **4**:1715–1723.
49. **Travers, A., and A. Klug.** 1987. DNA wrapping and writhing. Nature (London) **327**:280–281.
49a. **Walsh, K., and P. Schimmel.** 1987. Two nuclear factors compete for the muscle actin promoter. J. Biol. Chem. **262**:9429–9432.
50. **Weinberger, J., D. Baltimore, and P. A. Sharp.** 1986. Distinct factors bind to apparently homologous sequences in the immunoglobulin heavy-chain enhancer. Nature (London) **322**:846–848.
51. **Wright, S., A. Rosenthal, R. Flavell, and F. Grosveld.** 1984. DNA sequences required for regulated expression of β-globin genes in murine erythroleukemia cells. Cell **38**:265–273.
52. **Wu, B. J., R. E. Kingston, and R. I. Morimoto.** 1986. The human HSP70 promoter contains at least two regulatory domains. Proc. Natl. Acad. Sci. USA **83**:629–633.