A Novel Developmental Regulatory Motif Required for Stage-Specific Activation of the ϵ -Globin Gene and Nuclear Factor Binding in Embryonic Erythroid Cells

WILLIAM L. TREPICCHIO, MICHAEL A. DYER, AND MARGARET H. BARON*

*Department of Cellular and Developmental Biology, The Biological Laboratories,
Harvard University, Cambridge, Massachusetts 02138*

Received 1 November 1993/Returned for modification 14 February 1994/Accepted 8 March 1994

Members of the human β -globin gene family are expressed at discrete stages of development and therefore provide an important model system for examining mechanisms of temporal gene regulation. We have previously shown that expression of the embryonic β -like globin gene (ϵ) is mediated by a complex array of positive and negative upstream control elements. Correct developmental stage- and tissue-specific gene expression is conferred by synergistic interactions between a positive regulatory element (termed ϵ -PRE II) which is active only in embryonic erythroid cells and at least two other regulatory domains upstream of the ϵ -globin gene promoter. A nuclear factor highly enriched in cultured embryonic erythroid cells and in mouse embryonic yolk sac binds to a novel, evolutionarily conserved sequence within ϵ -PRE II. We show here that binding of this factor to the conserved element within ϵ -PRE II is critical for transcriptional activity. Point mutations that interfere with protein binding to ϵ -PRE II abolish transcriptional activation of the constitutive ϵ -globin promoter. Adult erythroid nuclei (from cultured cells or adult mouse liver) also contain a factor that binds to this region, but the complex formed migrates more rapidly during nondenaturing electrophoresis, suggesting either that distinct proteins bind to ϵ -PRE II or that a single protein is differentially modified in these cells in a way that modulates its activity. Several lines of evidence suggest that the binding factors in embryonic and adult erythroid cells are distinguished by posttranscriptional differences.

The human β -globin locus comprises a family of closely related genes which are expressed in a developmental stage-specific manner during erythroid cell ontogeny (28). The embryonic β -like globin gene (ϵ) is expressed early in gestation in the primitive erythroblasts of the yolk sac (12). Later in development, a change in gene expression (known as hemoglobin switching) occurs in which the embryonic gene is silenced and the fetal β -like globin genes are activated. Shortly after birth, a second switch occurs in which the fetal genes are down-regulated and the adult β -globin gene is turned on. Analysis of the coordinated expression of the β -globin locus has provided important insights into the developmental control of tissue-specific genes.

Hemoglobin switching events are regulated at several levels. Promoter elements that control constitutive expression of the embryonic, fetal, and adult globin genes have been characterized (see reference 30 and citations therein). In addition to promoter-proximal elements, a locus control region (LCR) is critical for high-level erythroid cell-specific expression of the entire β -globin gene family (6, 13). Competition between *cis*-acting, globin gene-proximal elements for stable interactions with the LCR is believed to be a central component of the fetal-to-adult β -globin switch (3, 6, 8, 9, 18). In contrast, the ϵ -globin gene is autonomously regulated and does not appear to be affected by the presence or absence of the other β -globin genes (24, 26).

We have recently reported that the upstream region of the ϵ -globin gene is complex and contains multiple positive and negative control elements (30). Of particular interest is a

positive regulatory element, termed ϵ -PRE II, which functions only in embryonic erythroid cells and interacts synergistically with another positive regulatory domain (ϵ -PRE V) located further upstream. Neither region is active on its own in monomeric or multimerized form, but in combination they cooperate to confer tissue- and developmental stage-specific expression on a constitutive ϵ -globin or heterologous promoter. The erythroid cell specificity but not the stage specificity of ϵ -PRE II is reconstituted in the presence of another element located further downstream. We have speculated that ϵ -PRE II and its associated proteins play an important role in the correct temporal expression of the ϵ -globin gene during development.

A nuclear factor that binds specifically to a novel sequence within ϵ -PRE II has been identified in extracts from embryonic erythroid cells (30). We show here that this element is broadly conserved throughout mammalian evolution and in simian primates contains a 6-bp dyad repeat. Clustered point mutations spanning this sequence specifically disrupt protein binding to ϵ -PRE II. The same mutations within ϵ -PRE II abolish transcriptional activation of the constitutive ϵ -globin promoter. Therefore, the DNA-binding activity of this factor is required for the regulatory function of ϵ -PRE II.

ϵ -PRE II-binding activity is found in nuclear extracts from adult as well as embryonic erythroid cell lines. Interestingly, the protein-DNA complexes formed by factors from these cells display different electrophoretic mobilities, suggesting that distinct proteins bind to ϵ -PRE II or that the same protein is differentially modified in the embryonic and adult erythroid cells (30). The relative differences in DNA-protein complex mobility observed for the erythroid cell lines are retained in normal embryonic and adult erythroid tissues and may reflect posttranscriptional modification of the same gene product in these cell types. We speculate that such a modified protein bound to ϵ -PRE II may be transcriptionally inactive in adult

* Corresponding author. Mailing address: Department of Cellular and Developmental Biology, Harvard University, The Biological Laboratories, 16 Divinity Ave., Cambridge, MA 02138. Phone: (617) 495-1135. Fax: (617) 495-9300. Electronic mail address: mbaron@husc4.harvard.edu.

erythroid cells. The ϵ -PRE II-binding factor may be a component of the temporal information that specifies accurate timing of globin gene switching during development.

MATERIALS AND METHODS

Mutation of ϵ -PRE II. PCR was used to introduce mutations within the dyad repeat of ϵ -PRE II (positions -438 to -423) or within a randomly chosen upstream region of ϵ -PRE II (positions -498 to -485), in the context of an amplified DNA fragment containing ϵ -globin sequences between -499 and -419. DNA fragments were amplified by using different pairs of wild-type (WT) and/or mutated oligonucleotides, the sequences of which are shown in Fig. 2. The fragment mutated within the dyad repeat region was designated MAD 19, and the fragment mutated within more upstream sequences of ϵ -PRE II was designated MAD 20. Double-stranded DNA sequencing was performed to confirm the sequences of the amplified products (25). Synthetic oligonucleotides used in some binding studies were the WT oligonucleotide 4 described in reference 30 and MAD 25, mutated at the same sites as the PCR oligonucleotide (see Fig. 2) but otherwise identical to oligonucleotide 4.

Plasmid constructions. LCR ϵ (-179)/CAT (30) contains the constitutive ϵ -globin promoter (-179 to +44) ligated to a chloramphenicol acetyltransferase (CAT) gene. This plasmid also contains a 2.5-kb μ LCR (10). ϵ -Globin regulatory regions, from positions -499 to -419 (ϵ -PRE II), -770 to -682 (ϵ -PRE V), and -770 to -419 (ϵ -PRE II to V), were isolated by PCR amplification as described previously (30). Sequences from -499 to -419 (ϵ -PRE II) and -770 to -682 (ϵ -PRE V) were subcloned in the native orientation into *Bgl*II-digested LCR ϵ (-179)/CAT. The resulting plasmid was designated construct 1. A larger fragment extending from -419 to -770 (containing ϵ -PRE II, ϵ -PRE V, and the intervening sequences) was also subcloned in the native orientation into the *Bgl*II site of LCR ϵ (-179)/CAT, to produce construct 4. Mutated PCR fragments (see above) were cloned in tandem with a WT ϵ -PRE V fragment in the correct orientation by ligation into *Bgl*II-digested LCR ϵ (-179)/CAT and were designated constructs 2 (mutations within ϵ -PRE II) and 3 (mutations upstream of ϵ -PRE II) (see Fig. 4). A larger (mutated) -419 to -770 fragment was subcloned into LCR ϵ (-179)/CAT and designated construct 5.

Cell culture, transfections, and enzymatic assays. K562, GM979, and MEL cells were maintained in culture as previously described (30). K562 cells were transfected by a modification of the DEAE-dextran method as previously described (30). β -Galactosidase activity (27) was used to correct for differences in transfection efficiency among samples. Constructs were tested in duplicate with two different DNA preparations in each of three transfection experiments. Enzymatic CAT assays and analyses were performed as described previously (30).

Nuclear extracts, EMSA, and partial proteolytic digestion. Nuclear extracts were prepared from 3 liters of erythroid cells by the method of Dignam et al. (5) as described previously (30). Extracts prepared from mouse yolk sac and adult liver were prepared as described previously (30), with the following modifications. Tissues were dissected from CD-1 mice (Charles River Laboratories) at day 12.5 postcoitus and were disrupted in 15 ml of phosphate-buffered saline at 0°C with 30 to 40 strokes of a Dounce homogenizer (pestle B). Membranes and cell debris were pelleted by centrifugation for 5 min at 500 rpm in a Sorvall J6B centrifuge and were subsequently discarded. Nuclei were isolated and extracted from dissociated yolk sac and liver cells as described previously for cultured cells (30). The binding reaction conditions and nondenaturing gel

FIG. 1. Stage-specific regulation of the human ϵ -globin gene. (A) General organization of the human ϵ -globin upstream control region and model for positive regulation of the ϵ -globin gene. The 5' and 3' boundaries of four positive regulatory elements (ϵ -PRE I, II, IV, and V) that are active in embryonic erythroid cells (30) are indicated by black boxes. Three negative regulatory elements (ϵ -NRE I, II, and III) are indicated by hatched boxes, and the constitutive ϵ -globin promoter is indicated by a white box. Curved arrows represent synergistic interactions between ϵ -PRE II and ϵ -PRE V (which confer developmental stage- and erythroid cell-specific gene expression) and between ϵ -PRE II and ϵ -PRE I (which confer erythroid cell-specific gene expression) (30). The figure, which was modified from Fig. 5 of reference 30, is not drawn to scale. (B) Dyad repeat within ϵ -PRE II and phylogenetic comparison of the sequence. A portion of the sequence (31) of ϵ -PRE II analyzed by methylation interference (30) is shown at the top. The 6-bp dyad repeat of two 5'-CTGTT(A/T)-3' half-sites is indicated by arrows. A phylogenetic comparison of ϵ -PRE II sequences revealed a 14-bp evolutionarily conserved sequence, shown at the bottom. Gibbon and orangutan sequences were obtained from a data base comparison with the human sequence. The relevant mouse, rabbit, and galago sequences used for this comparison were taken from reference 15. Bases conserved throughout evolution are indicated by asterisks.

system have been described previously (30). The probes for the electrophoretic mobility shift assays (EMSAs) were radiolabeled by filling in the *Bam*HI and *Bgl*II half-sites with [α - 32 P]dCTP, using Klenow DNA polymerase. Serial dilutions of proteinase K (EM Science) and trypsin (Sigma) were made in 50 mM Tris-HCl (pH 7.5)-5 mM CaCl_2 . One microliter of diluted protease was added to each binding reaction (30 min, room temperature; 8 μ g of protein, 0.6 ng of probe). Proteolysis was allowed to proceed at room temperature for 5 min, and samples were then loaded onto a nondenaturing acrylamide gel. The octamer oligonucleotide used as a control in this experiment has been described previously (30).

Gel filtration chromatography. Nuclear extract (2.5 mg) prepared from GM979 or MEL cells was chromatographed on a column (1.0 by 16 cm) of Sepharose CL-6B in buffer D (5). Fractions were collected, and aliquots were assayed for ϵ -PRE II-binding activity by EMSA. The column was calibrated by using a panel of typical globular proteins.

FIG. 2. Nucleotide sequence of the ϵ -globin gene between -499 and -419 and oligonucleotide primers for PCR amplification. The nucleotide sequence of the 80-bp PCR fragment containing ϵ -PRE II is shown. This fragment, when ligated in tandem with ϵ -PRE V, confers developmental stage- and tissue-specific gene expression on the constitutive ϵ -globin promoter (30). The ϵ -PRE II regulatory element defined by deletional analysis is indicated by a hatched bar. Horizontal arrows denote the 6-bp dyad repeat within ϵ -PRE II. Guanosine residues involved in critical protein contacts are indicated by open circles. Wild-type and mutated (mut) 5' (sense) and 3' (antisense) oligonucleotide primer sequences are presented, respectively, above and below the sequence of the PCR fragment. Oligonucleotides were synthesized with *Bgl*II (5' primers) or *Bam*HI (3' primers) restriction endonuclease sites (indicated by brackets and lowercase letters) to facilitate cloning and multimerization of the final PCR products. Black dots and vertical arrows represent mutated nucleotides.

Kinetics of association and dissociation of complexes A and B. Association kinetics were examined for binding reactions containing binding buffer (30), oligonucleotide 4 (30), and poly(dI-dC) (3 μ g) in 18 μ l. Reactions were equilibrated at room temperature for 5 min. Nuclear extract (2 μ l of 2 mg/ml) was added at times 0, 10 min, 20 min, 22.5 min, 25 min, 27.5 min, and 29 min, and all of the reactions were loaded onto a nondenaturing gel at time 30 min. Dissociation kinetics were examined by incubating 18- μ l binding reactions at room temperature for 30 min. A 100-fold excess of unlabeled probe (oligonucleotide 4, containing the dyad repeat) was added, and the samples were loaded onto a nondenaturing gel at the times indicated. Band intensities were determined by densitometry of preflashed autoradiograms; the resulting data were used to estimate a $t_{1/2}$ for the dissociation reaction.

RESULTS

An evolutionarily conserved element is critical for embryonic erythroid nuclear factor binding to ϵ -PRE II. We have recently described a cluster of multiple positive and negative *cis*-acting control elements upstream of the constitutive promoter of the human embryonic β -like globin gene (ϵ) (30). One of these positive elements, termed ϵ -PRE II, is of particular interest because it confers developmental stage-specific expression on the ϵ -globin promoter through synergistic interactions with other ϵ -globin regulatory domains (summarized in Fig. 1A). A nuclear factor highly enriched in human and mouse embryonic erythroid cells binds to a 19-bp region of ϵ -PRE II (30). This sequence contains a novel dyad repeat of two 5'-CTGTT(T/A)-3' half-sites that could potentially bind a dimeric protein (30). Methylation interference analysis identified three guanosine residues involved in essential protein contacts with ϵ -PRE II. Two of these residues fall within the dyad repeat (30) (Fig. 1B, top).

A number of regions upstream of the ϵ -globin gene are phylogenetically conserved (14, 15). One of the conserved regions noted by Hardison and colleagues (15) overlaps the sequence that we have identified as ϵ -PRE II (30). As shown in Fig. 1B, a 14-bp region within ϵ -PRE II is broadly conserved throughout mammalian evolution. The dyad repeat sequence

of ϵ -PRE II is perfectly conserved in the simian primates orangutans and gibbons (Fig. 1B), which, like humans, undergo two β -globin switches during development (16, 23). Although the dyad repeat per se is not well conserved in mouse, rabbit, or galago sequences, nuclear factors from both mouse and human erythroid cells form specific complexes with oligonucleotides containing either the human (30), mouse, or rabbit sequences (not shown).

To confirm that the conserved element within ϵ -PRE II plays a significant role in protein binding, point mutations were introduced into the putative recognition sequence. In choosing nucleotides for mutation, we were guided in part by the results of the methylation interference experiments (30). A set of six clustered point mutations was introduced into a PCR oligonucleotide primer (MAD 19) across the region containing the dyad repeat (Fig. 2). As a control, a cluster of six random point mutations was introduced into the sequence upstream of the dyad repeat, through an oligonucleotide primer denoted MAD 20 (Fig. 2). Under the conditions used for these EMSA analyses, we have not detected any specific protein binding to the region mutated in MAD 20 (7). The mutated oligonucleotides and their corresponding WT controls were then used to amplify by PCR the 80-bp ϵ -globin sequence from positions -419 to -499. By including different primer pairs in the PCRs, three fragments, containing the WT sequence, mutations across the dyad repeat (MAD 19), and mutations upstream of the dyad repeat, between positions -499 and -486 (MAD 20), were generated (Fig. 2).

The PCR-amplified fragments were analyzed for protein binding by EMSA (Fig. 3A). Specific complexes were formed by nuclear factors from embryonic (mouse GM979 [lane 1] and human K562 [not shown]) and adult (MEL) (lane 3) mouse erythroid cells. The PCR product-containing complexes comigrated with complexes formed with synthetic double-stranded oligonucleotides used in our previous studies (30) (data not shown). The complexes formed by nuclear factors from embryonic (GM979, complex A) and adult (MEL, complex B) erythroid cells migrated with different mobilities (Fig. 3A; compare lanes 1 and 3), consistent with our previous observations (30). Mutations within the dyad repeat region of ϵ -PRE

FIG. 3. Protein binding to WT and mutated ϵ -PRE II fragments. (A) PCR probes. Factors from embryonic (GM979; lane 1) and adult (MEL; lane 3) erythroid nuclei form sequence-specific complexes (A and B, respectively) with the WT PCR fragment containing ϵ -PRE II (see Fig. 2 for sequence). A PCR fragment (MAD 19) mutated across the dyad repeat of ϵ -PRE II failed to form complex A (lane 6) or B (lane 8). Introduction of a cluster of mutations within the sequence upstream of ϵ -PRE II had no effect on formation of either complex A (lane 11) or complex B (lane 13). Formation of complexes A and B was specifically competed for when a 20-fold molar excess of unlabeled double-stranded synthetic ϵ -PRE II (30) oligonucleotide was included in the binding reaction (lanes 2 and 4). (Competition for complex formation was significantly more efficient when an unlabeled, double-stranded synthetic ϵ -PRE II oligonucleotide was used instead of the unlabeled purified PCR product.) The three 80-bp PCR fragments used in the EMSAs are diagrammed at the bottom. Arrows indicate the position of the dyad repeat within the WT fragment; X's denote clustered point mutations. -NE, no nuclear extract added to the binding reaction. (B) Synthetic oligonucleotides. For these binding reactions, synthetic oligonucleotides were the WT oligonucleotide 4 from reference 30 (here termed MAD 4) and MAD 25, mutated at the same sites as the PCR oligonucleotide (Fig. 2) but otherwise identical to oligonucleotide 4. Binding reactions were carried out with nuclear extracts from embryonic erythroid (GM979 and K562; complex A) or adult erythroid (MEL; complex B) cells, in the absence (-; lanes 1, 3, 5, 7, 8, 10, 12, and 14) or presence (+) of a 20-fold excess of unlabeled competitor oligonucleotide (unlabeled MAD 4 [lanes 2, 4, and 6] or unlabeled MAD 25 [lanes 9, 11, and 13]). The synthetic oligonucleotides used in the EMSAs are diagrammed at the bottom. Arrows indicate the position of the dyad repeat within the WT fragment; X's denote clustered point mutations. -NE, no nuclear extract added to the binding reaction.

II disrupted binding of proteins from both embryonic (lane 6) and adult (lane 8) erythroid cells. In contrast, mutations within the sequence upstream of the dyad repeat had no effect on binding (Fig. 3A, lanes 11 to 14). Binding specificity was determined by competition with an unlabeled double-stranded synthetic ϵ -PRE II oligonucleotide (lanes marked +). Protein from embryonic (GM979 and K562) and adult erythroid (MEL) nuclear extracts also formed specific complexes (Fig. 3B, lanes 1, 3, and 5) with a shorter WT synthetic oligonucleotide (oligonucleotide 4 [30]) but not with a synthetic oligonucleotide (MAD 25) containing the same mutations as PCR probe MAD 19 (Fig. 3B, lanes 8, 10, and 12). We conclude from these results that the conserved 14-bp element within ϵ -PRE II is part of the recognition sequence of the ϵ -PRE II-binding factor.

Mutations that interfere with protein binding to the dyad repeat also abolish transcriptional activation by ϵ -PRE II. To correlate protein binding to the dyad repeat with transcrip-

tional activation by ϵ -PRE II, the WT and mutated fragments used in the EMSA analyses described above were subcloned upstream of the ϵ -globin constitutive promoter (-179 to +44) in LCR ϵ (-179)/CAT (Fig. 4) (30). The μ LCR is required for ϵ -promoter-driven CAT activity under our assay conditions (30). We have previously shown that ϵ -PRE II and another upstream regulatory domain (ϵ -PRE V) are inactive in isolation but interact synergistically to confer developmental stage- and tissue-specific gene expression in a transient assay system (30) (Fig. 1A). Therefore, a fragment containing ϵ -PRE V (Fig. 4) was ligated in tandem with either the wild-type ϵ -PRE II or the mutation-containing fragments.

Analysis of these constructs in embryonic erythroid cells (K562) demonstrated that mutations within ϵ -PRE II which disrupt protein binding also abolish transcriptional activation (Fig. 4). The constructs containing the tandemly ligated ϵ -PRE II and ϵ -PRE V fragments (Fig. 4, constructs 1 and 2) and constructs containing a larger regulatory region (30) with

FIG. 4. Functional analysis of WT and mutated ϵ -PRE II fragments. The ϵ -globin upstream regulatory region is shown schematically at the top. ϵ -PRE II and ϵ -PRE V are indicated by black boxes, and the constitutive ϵ -globin promoter is indicated by a white box. The PCR fragments represented in lines A and B were generated as described in Materials and Methods. The WT and mutated ϵ -PRE II PCR fragments (for sequences, see Fig. 2) are represented as a solid black box in line A. The hatched box (line A) represents a previously described 85-bp PCR fragment containing ϵ -PRE V (30). The shaded box (line B) represents a PCR fragment containing sequences from -770 to -419 (ϵ -PRE V through ϵ -PRE II) (30). ϵ -PRE II (WT or mutated) and ϵ -PRE V (line A) fragments were subcloned in tandem in the correct orientation into constitutive ϵ -globin promoter constructs (constructs 1 to 3). Construct 1 contains the WT ϵ -PRE II sequence, construct 2 contains a cluster of six point mutations introduced within ϵ -PRE II (depicted by XX), construct 3 contains a cluster of six mutations introduced upstream of ϵ -PRE II (depicted by XX), and construct 4 contains a WT -770 to -419 fragment; in construct 5, point mutations (XX) were introduced into ϵ -PRE II (sequence shown in Fig. 2). All constructs contained the constitutive ϵ -globin promoter (represented by a small black box), ligated upstream from a CAT reporter gene, and the μ LCR (10). The constructs were analyzed in human embryonic erythroid (K562) and mouse adult erythroid (MEL) cells. Values graphed represent averages of three experiments $\pm$ standard deviations. Acetylation levels are expressed relative to the activity of the mutated ϵ -PRE II construct (construct 2, group A; construct 5, group B).

sequences between and including ϵ -PRE II (WT or mutated) and ϵ -PRE V (constructs 4 and 5) gave similar results. Mutations upstream of ϵ -PRE II (MAD 20) which did not affect protein binding also had no effect on transcriptional activity (Fig. 4, construct 3). Neither set of mutations had a significant effect in adult mouse erythroid (MEL) cells, consistent with our earlier observations that a fragment containing ϵ -PRE II is not functional in adult erythroid cells (30). Therefore, protein binding within the 14-bp conserved element is critical for developmental stage-specific transcriptional activation by ϵ -PRE II. A more detailed analysis involving single base changes within the dyad repeat will be required to determine the relative contributions of individual residues to binding and transcriptional activity.

Nuclear factors from mouse embryonic and adult erythroid tissues form distinct complexes with ϵ -PRE II. Embryonic and adult erythroid nuclei contain factors which make identical contacts with guanosine residues within ϵ -PRE II but form complexes of different mobilities (A and B, respectively) by nondenaturing acrylamide gel electrophoresis (30). Little or no binding activity is detected in the soluble fraction (S100) of the cell (7). These observations suggest either that ϵ -PRE II is bound by different proteins in embryonic and adult erythroid cells or that the same protein is differentially modified in these cell types (the latter category of possibilities includes different isoforms generated by alternative splicing). The embryonic and adult erythroid cell lines used in these studies were derived from different mouse strains: DDD (GM979 cells [4]) and

FIG. 5. Binding analyses for embryonic and adult erythroid nuclear extracts. (A) EMSA was performed with a synthetic oligonucleotide containing the dyad repeat of ϵ -PRE II (oligonucleotide 4 of reference 30) and nuclear extracts prepared from embryonic (GM979, K562, and yolk sac; lanes 1 and 3 to 6) and adult (MEL and adult liver; lanes 2 and 7 to 9) erythroid cells. Embryonic erythroid factors bound to the ϵ -PRE II probe to form complex A. Factors from adult erythroid cells formed a more slowly migrating complex, B. A faster-migrating complex (unlabeled gray arrow) is observed in human K562 cells and has been previously shown to be nonspecific (30). Complex formation was inhibited when challenged with a 20-fold molar excess of unlabeled specific ϵ -PRE II oligonucleotide (30) (lanes 5 and 8) but was not affected by the addition of a nonspecific oligonucleotide containing an octamer motif (lanes 6 and 9). -NE, no nuclear extract added to the binding reaction; F, free (unbound) probe. (B) Embryonic and adult erythroid nuclear extracts from cell lines and normal mouse tissues form complexes with distinct mobilities. EMSA was performed with a synthetic ϵ -PRE II oligonucleotide as described above. The resolution between complexes A and B was improved by running the gel at 200 V for 8 h (instead of the usual 4 h). Complexes formed by factors in embryonic erythroid (GM979, yolk sac; complex A) and adult erythroid (MEL, adult liver; complex B) nuclear extracts show clearly distinct mobility differences (compare lanes 1 and 2 and lanes 3 and 4).

DBA/2 (MEL [11]), respectively. To rule out the possibility that the observed mobility differences were mouse strain dependent or an artifact of the cultured cell system, nuclear extracts were prepared from the adult liver (a source of normal adult erythroid cells) and yolk sac tissues (a source of normal embryonic erythroid cells) of a pregnant CD-1 mouse (day 12.5 postcoitus).

Yolk sac and adult liver nuclei contain factors that bind to ϵ -PRE II to form complexes which are indistinguishable from the corresponding complexes formed by factors from GM979 or MEL cell nuclei (Fig. 5A; compare lanes 1 and 4 and lanes 2 and 7). The mobility differences observed previously for embryonic and adult erythroid factors in cultured cells (Fig. 5A [compare lanes 1 and 3 with lane 2] and 5B [compare lanes 1 and 2]) are retained in extracts from normal mouse tissues (Fig. 5A [compare lanes 1, 3, and 5 and lanes 2 and 7] and 5B [compare lanes 3 and 4]). These mobility differences are

FIG. 6. Limited proteolysis suggests that the proteins in complexes A and B are very similar. The results of digestion with proteinase K are shown. The amounts of protease used, in micrograms per 20- μ l reaction, were 0 (lanes 3, 8, and 13), 0.01 (lanes 4, 9, and 14), 0.1 (lanes 5, 10, and 15), 1.0 (lanes 6, 11, and 16); and 10.0 (lanes 7, 12, and 17). Eight micrograms of nuclear extract was incubated with 0.6 ng of labeled probe (oligonucleotide 4 or an octanucleotide [octa] control [30]). Protease digestions were carried out for 5 min at room temperature, and the reaction contents were loaded onto a nondenaturing gel for EMSA. The smaller digestion products from the ϵ -PRE II-binding complex that appear to comigrate with products from the octamer (Oct)-binding protein digestion are nonspecific and do not compete with unlabeled oligonucleotide (not shown).

consistent among three different strains of mice (DBA/2, DDD, and the outbred strain CD-1) and therefore appear to be independent of genetic background.

Comparison of complexes A and B by partial proteolytic digestion and gel filtration chromatography. As a first step toward determining whether embryonic and adult erythroid nuclei contain distinct ϵ -PRE II-binding activities, or if the same protein might be differentially modified in the two cell types, limited proteolytic digestion of ϵ -PRE II-specific complexes formed by factors from embryonic and adult erythroid nuclear extracts was performed. Partial proteolysis often produces subspecies which retain their specific DNA-binding activities, and different proteins generate distinct cleavage products that can be monitored by EMSA (20, 22).

Virtually identical patterns of proteolytic digestion were obtained for the complexes formed by proteins in GM979 and MEL extracts (Fig. 6). In contrast, the digestion pattern obtained for an octamer-binding complex in GM979 extracts (presumably reflecting Oct-1 activity [29]) was clearly different (compare lanes 3 to 7 and lanes 13 to 17). Partial proteolysis produced intermediate products of the same apparent mobility for ϵ -PRE II-binding factors from both extracts (compare lanes 9 to 12 and lanes 14 to 17). Similar results were obtained when the extracts were treated with trypsin (data not shown), suggesting that the protein components of the complexes formed by factors from the embryonic and adult erythroid cells are very similar and possibly identical.

FIG. 7. Kinetics of association and dissociation of complexes A and B. (A) Time course of association. Binding reactions containing binding buffer (30), oligonucleotide 4 (see reference 30 and Materials and Methods), and dI-dC (3 μ g) in 18 μ l were equilibrated at room temperature for 5 min. Nuclear extract (2 μ l of 2 mg/ml) was added at times 0, 10 min, 20 min, 22.5 min, 25 min, 27.5 min, and 29 min, and all of the reactions were loaded onto a nondenaturing gel at time 30 min. The incubation time is indicated above each lane. (B) Time course of dissociation. Binding reactions containing binding buffer (30), oligonucleotide 4 (see reference 30 and Materials and Methods), and poly(dI-dC) (3 μ g) in 18 μ l were incubated at room temperature for 30 min. Under these conditions, 30 min was sufficient to equilibrate the binding reaction (see panel A). A 100-fold excess of unlabeled probe (oligonucleotide 4, containing the dyad repeat) was added, and the samples were loaded onto a nondenaturing gel at the time indicated above each lane. To ensure that the decrease in binding activity was due to competition by unlabeled probe rather than to protein instability, a binding reaction was incubated for the longest time point (a total of 75 min at room temperature) without unlabeled competitor (none; lanes 5 and 10). -P, no protein added; F, free probe.

The molecular mass of the ϵ -PRE II-binding factor from embryonic (GM979) and adult (MEL) erythroid cells was estimated by gel filtration chromatography on Sepharose CL-6B. Identical elution volumes were obtained for the ϵ -PRE II-binding activities from GM979 and MEL cell nuclei, yielding an estimated molecular mass of 80 to 85 kDa.

Kinetics of protein binding to ϵ -PRE II. The apparent ease of detection of specific protein binding to ϵ -PRE II by EMSA is consistent either with high-affinity binding or high factor abundance or both. An indication of binding affinity or stability would be useful in guiding the biochemical purification of the ϵ -PRE II-binding activity. We therefore directly examined the kinetics of nuclear factor binding to an ϵ -PRE II probe. Both complexes A (GM979 embryonic erythroid) and B (MEL adult erythroid) formed very rapidly and were easily detected within 1 min of incubation (Fig. 7A). The amount of bound complex A or B formed in an association reaction at room temperature was maximal by about 10 min and remained stable for at least 30 to 60 min (Fig. 7 and data not shown). A 60-min incubation at 37°C had no detectable effect on the stability of the protein- ϵ -PRE II complex (data not shown).

For the dissociation rate determinations (Fig. 7B), the DNA-protein binding reactions were allowed to reach equilibrium (about 30 min; Fig. 7A). A 100-fold molar excess of unlabeled specific competitor was added, and aliquots were loaded onto the running gel at various times after competitor addition (lanes 1 to 4 and 6 to 9; no competitor was added to the samples run in lanes 5 and 10). Each complex, A and B, once formed, was extremely stable; whereas a 10-fold molar

excess of unlabeled specific oligonucleotide interfered completely with binding, once the complex had formed, even a 100-fold molar excess reduced but did not eliminate binding. The $t_{1/2}$ for the dissociation reaction was estimated as ≥ 45 min (see Materials and Methods).

DISCUSSION

Our analysis of human ϵ -globin gene expression in cultured cells has allowed us to identify components of what appears to be a complex system of combinatorial interactions (30). At least two positive regulatory elements (ϵ -PRE II and ϵ -PRE V) interact synergistically to direct developmentally appropriate expression of this gene in embryonic erythroid cell lines. We have previously identified a nuclear factor highly enriched in human and mouse embryonic erythroid cells that binds to a novel sequence motif within ϵ -PRE II (30). Although ϵ -PRE II-binding activity is found in adult erythroid cells (reference 30 and data presented here), it migrates with somewhat faster mobility by nondenaturing gel electrophoresis. No other factors have been found by EMSA to bind to ϵ -PRE II sequences upstream from the dyad repeat (7). The ϵ -PRE II-binding factor in embryonic erythroid cells is a candidate regulator of developmental stage-specific ϵ -globin gene expression. Because of the potentially central role of ϵ -PRE II and its associated proteins in the activation of the ϵ -globin gene during embryogenesis, the significance of the observed protein binding was examined further.

We show here that a 14-bp region of ϵ -PRE II is broadly

conserved throughout mammalian evolution, consistent with its proposed role in developmental regulation of the ϵ -globin gene. Point mutational analysis of this region demonstrates that this element is critical for ϵ -PRE II binding by factors in human as well as mouse erythroid nuclei. Protein binding correlates well with transcriptional activity: mutations that disrupt binding also abolish activation of the ϵ -globin promoter by ϵ -PRE II.

It is not understood how ϵ -PRE II cooperates with ϵ -PRE V to direct stage-specific activation. ϵ -PRE V is required for the stage-specific activity of ϵ -PRE II in erythroid cells (30) and may play a role in defining the activity of ϵ -PRE II at different stages of development. The region containing ϵ -PRE V is complex, by at least two criteria: (i) it interacts synergistically with ϵ -PRE II to faithfully reconstitute tissue- and stage-specific expression, and (ii) it activates expression in embryonic erythroid cells and suppresses expression in adult erythroid cells (30). ϵ -PRE V may alternate during development between a positive or negative control function or may contain more than one distinct regulatory element (see reference 30 for a discussion). We have undertaken an analysis of protein binding to ϵ -PRE V. Preliminary results (17) reveal a complex organization of protein-binding sites within ϵ -PRE V, consistent with the functional complexity revealed by expression studies in cultured cells (30).

Analysis of globin gene activation in transient heterokaryons has demonstrated that regulation of embryonic β -like globin genes is mediated by stage-specific *trans*-acting molecules (1, 2). The ϵ -PRE II-binding factor that we have identified may be one such stage-specific regulator. Paradoxically, ϵ -PRE II-binding activity has been found in extracts from adult erythroid cells (reference 30 and this report), in which ϵ -PRE II is inactive (30). However, the factors from embryonic and adult erythroid cells form complexes of different electrophoretic mobilities (30). The relative differences in mobility observed for complexes formed by factors from embryonic and adult erythroid cell lines are retained in normal liver and yolk sac tissues from the same pregnant mouse. The cell lines and the mouse tissues were from three different strains of mouse, suggesting that the mobility differences are independent of genetic background and may be of functional significance.

Although the factors responsible for ϵ -PRE II-specific binding in embryonic and adult erythroid nuclei may be different proteins, two major lines of evidence suggest that the differences in the complexes that they form reflect posttranscriptional modifications. The methylation interference patterns are identical for embryonic and adult erythroid binding activities, as are their partial proteolytic digestion products. Equilibrium DNA-protein binding experiments indicate that the kinetics of association and dissociation are very similar for both the embryonic and adult erythroid factors, providing additional (if less direct) evidence for the similarity of these binding activities. Finally, the elution behaviors of the two binding activities by gel filtration chromatography are indistinguishable.

We speculate that the inactivity of ϵ -PRE II in adult erythroid cells may be due simply to differential, stage-specific modification of the binding factor. Inactivation of the ϵ -PRE II-binding factor by posttranscriptional modification may also be a component of the down-regulation of ϵ -globin gene expression observed at later stages of development. Purification and molecular cloning of the ϵ -PRE II-binding factor will be required to resolve this question. The relatively slow rate of dissociation of the ϵ -PRE II-binding factor from its cognate site ($t_{1/2}$ of ≥ 45 min) suggests that DNA affinity chromatography (19, 21) may provide a very efficient purification step for this protein.

ACKNOWLEDGMENTS

W.L.T. and M.A.D. contributed equally to this work.

We are grateful to James Bieker and to members of our laboratory for thoughtful reading of the manuscript. Special thanks go to Frank Conlon for assistance with mouse tissue dissections and to Patrick Hayes for oligonucleotide synthesis.

This work was supported by grants (to M.H.B.) from the National Institutes of Health (RO1 GM42413), the Lucille P. Markey Charitable Trust (grant 87-24), and the March of Dimes Birth Defects Foundation (Basil O'Connor Award 5-715). W.L.T. is a postdoctoral fellow of the Cooley's Anemia Foundation. M.A.D. is supported by NIH predoctoral training grant GM 07598. M.H.B. is a Lucille P. Markey Scholar in Biomedical Science.

REFERENCES

1. Baron, M. H., and T. Maniatis. 1986. Rapid reprogramming of globin gene expression in transient heterokaryons. *Cell* **46**:591-602.
2. Baron, M. H., and T. Maniatis. 1991. Regulated expression of human α - and β -globin genes in transient heterokaryons. *Mol. Cell. Biol.* **11**:1239-1247.
3. Behringer, R. R., T. M. Ryan, R. D. Palmiter, R. L. Brinster, and T. M. Townes. 1990. Human gamma- to beta-globin switching in transgenic mice. *Genes Dev.* **4**:380-389.
4. Brown, B. A., R. W. Padgett, S. C. Hardies, C. A. Hutchison, and M. H. Edgell. 1982. β -Globin transcript found in induced murine erythroleukemia cells is homologous to the β h0 and β h1 genes. *Proc. Natl. Acad. Sci. USA* **79**:2753-2757.
5. 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.
6. Dillon, N., and F. Grosfeld. 1993. Transcriptional regulation of multigene loci: multilevel control. *Trends Genet.* **9**:134-137.
7. Dyer, M. A., and M. H. Baron. Unpublished data.
8. Engel, J. D. 1993. Developmental regulation of human β -globin gene transcription: a switch of loyalties? *Trends Genet.* **9**:304-309.
9. Enver, T., N. Raich, A. J. Ebens, T. Papayannopoulou, F. Costantini, and G. Stamatoyannopoulos. 1990. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature (London)* **344**:309-313.
10. Forrester, W., U. Novak, R. Gelinas, and M. Groudine. 1989. Molecular analysis of the human beta-globin locus activation region. *Proc. Natl. Acad. Sci. USA* **86**:5439-5443.
11. Friend, C., W. Scher, J. G. Holland, and T. Sato. 1971. Hemoglobin synthesis in murine virus-induced leukemic cells in vitro: stimulation of erythroid differentiation by dimethyl sulfoxide. *Proc. Natl. Acad. Sci. USA* **68**:378-382.
12. Gale, R. E., J. B. Clegg, and E. R. Huehns. 1979. Human embryonic haemoglobins Gower 1 and Gower 2. *Nature (London)* **280**:162-164.
13. Grosfeld, F., G. B. van Assendelft, D. Greaves, and G. Kollias. 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* **51**:975-985.
14. Gumucio, D. L., D. A. Shelton, W. J. Bailey, K. Hayasaka, J. L. Slightom, and M. Goodman. 1993. Phylogenetic footprinting reveals unexpected complexity in trans factor binding upstream from the ϵ globin gene. *Proc. Natl. Acad. Sci. USA* **90**:6018-6022.
15. Hardison, R., K.-M. Chao, M. Adamkiewicz, D. Price, J. Jackson, T. Zeigler, N. Stojanovic, and W. Miller. 1993. Positive and negative regulatory elements of the rabbit embryonic ϵ -globin gene revealed by an improved multiple alignment program and functional analysis. *DNA Sequence* **4**:163-176.
16. Hardison, R. C. 1990. Evolution of globin gene families, p. 272-289. *In* R. K. Selander, A. G. Clark, and T. S. Whittam (ed.), *Evolution at the molecular level*. Sinauer Associates, Inc., Sunderland, Mass.
17. Hirose, K., W. Trepicchio, and M. Baron. Unpublished data.
18. Jane, S. M., P. A. Ney, E. F. Vanin, D. L. Gumucio, and A. W. Nienhuis. 1992. Identification of a stage selector element in the human γ -globin gene promoter that fosters preferential interaction with the 5' HS2 enhancer when in competition with the β -promoter. *EMBO J.* **11**:2961-2969.
19. Kadonaga, J. T. 1991. Purification of sequence-specific binding proteins by DNA affinity chromatography. *Methods Enzymol.* **208**:10-23.

20. **Karnitz, L., D. Poon, P. A. Weil, and R. Chalkley.** 1989. Purification and properties of the Rous sarcoma virus internal enhancer binding factor. *Mol. Cell. Biol.* **9**:1929–1939.
21. **Larson, C. J., and G. L. Verdine.** 1992. A high-capacity column for affinity purification of sequence-specific DNA-binding proteins. *Nucleic Acids Res.* **20**:3525.
22. **Lee, T. C., and R. J. Schwartz.** 1992. Using proteases to avoid false identification of DNA-protein complexes in gel shift assays. *Bio-Techniques* **12**:486–490.
23. **Li, W.-H., M. Gouy, P. Sharp, C. O'Huigin, and Y.-W. Yang.** 1990. Molecular phylogeny of rodentia, lagomorpha, primates, artiodactyla and carnivora and molecular clocks. *Proc. Natl. Acad. Sci. USA* **87**:6703–6707.
24. **Raich, N., T. Enver, B. Nakamoto, B. Josephson, T. Papayannopoulou, and G. Stamatoyannopoulos.** 1990. Autonomous developmental control of human embryonic globin gene switching in transgenic mice. *Science* **250**:1147–1149.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
26. **Shih, D. M., R. J. Wall, and S. G. Shapiro.** 1990. Developmentally regulated and erythroid-specific expression of the human embryonic β -globin gene in transgenic mice. *Nucleic Acids Res.* **18**:5465–5472.
27. **Simon, J., and J. Lis.** 1987. A germline transformation analysis reveals flexibility in the organization of heat shock consensus elements. *Nucleic Acids Res.* **15**:2971–2988.
28. **Stamatoyannopoulos, G., and A. Nienhuis.** 1987. Hemoglobin switching, p. 66–105. *In* G. Stamatoyannopoulos, A. W. Nienhuis, P. Leder, and P. W. Majerus (ed.), *The molecular basis of blood diseases*. The W. B. Saunders Co., Philadelphia.
29. **Sturm, R. A., G. Das, and W. Herr.** 1988. The ubiquitous octamer binding protein Oct-1 contains a POU domain with a homeo subdomain. *Genes Dev.* **2**:1582–1599.
30. **Trepicchio, W. L., M. A. Dyer, and M. H. Baron.** 1993. Developmental regulation of the human embryonic β -globin gene is mediated by synergistic interactions among multiple tissue- and stage-specific elements. *Mol. Cell. Biol.* **13**:7457–7468.
31. **Trepicchio, W. L., M. A. Dyer, R. C. Hardison, and M. H. Baron.** Upstream regulatory region of the human embryonic β -like globin gene, ϵ . DNA Sequence, in press.