Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 7693–7697, October 1990
Genetics

Conservation of the primary structure, organization, and function of the human and mouse β-globin locus-activating regions

(β-globin genes/enhancers/dominant control region/DNA sequence homology/K562 erythroleukemia cells)

ANNE M. MOON AND TIMOTHY J. LEY*

Division of Hematology-Oncology, Departments of Medicine and Genetics, Jewish Hospital at Washington University Medical Center, 216 South Kingshighway Boulevard, Saint Louis, MO 63110

Communicated by Philip W. Majerus, July 16, 1990 (received for review June 15, 1990)

ABSTRACT **DNA sequences located in a region 6–18 kilobases (kb) upstream from the human ε-globin gene are known as the locus-activating region (LAR) or dominant control region. This region is thought to play a key role in chromatin organization of the β-like globin gene cluster during erythroid development. The β-globin LAR activates linked globin genes in transiently or stably transfected erythroleukemia cells and in erythroid cells of transgenic mice. Since the human β-globin LAR is functional in mice, we reasoned that critical LAR sequence elements might be conserved between mice and humans. We therefore cloned murine genomic sequences homologous to one portion of the human LAR (site II, positions −11,054 to −10,322 with respect to the human ε gene). We found that this murine DNA fragment (mouse LAR site II) and sequences homologous to human LAR sites I and III are located upstream from the mouse β-like globin gene cluster and determined that their locations relative to the cluster are similar to that of their human counterparts. The homologous site II sequences are 70% identical between mice and humans over a stretch of ≈800 base pairs. Multiple core sequences with >80% identity were present within this region. Transient and stable transfection assays of K562 erythroleukemia cells demonstrated that both human and mouse LAR elements contain enhancer activity and confer hemin inducibility on a linked human γ-globin promoter. These results suggest that primary structural elements—and the spatial organization of these elements—are important for function of the β-globin LAR.**

The human β-globin locus-activating region (hLAR) was first described as a series of erythroid-specific, developmentally stable DNase I-hypersensitive sites located upstream and downstream of the β-like globin gene cluster (1, 2). hLAR site I is located at position −6.1 kilobases (kb); site II, at −10.9 kb; site III, at −14.7 kb; and site IV, at −18 kb with respect to the cap site of the human ε-globin gene. These hypersensitive sites are reconstituted when chromosome 11 from nonerythroid cells is introduced into mouse erythroleukemia (MEL) cells, suggesting that the locus itself contains all of the cis-acting DNA elements necessary for chromatin reorganization (3).

Several groups have observed that LAR sequences are required for full activity of the human α- or β-globin genes in transgenic mice (4–10). Further studies of the hLAR in transgenic mice indicated that site II alone contains at least 50% of the "activity" of all four sites together (5, 6, 10). hLAR function has also been studied with *in vitro* expression assays. Tuan *et al.* (11) added individual hLAR sites to a plasmid containing an ε-globin promoter driving the chloramphenicol acetyltransferase gene (ε-CAT) and transfected these plasmids into K562, MEL, or nonerythroid cells. Only hLAR site II enhanced expression of ε-CAT; this effect was erythroid-specific. Similarly, a human β-globin gene, the Thy-1 gene, or a thymidine kinase promoter-driven neomycin phosphotransferase hybrid gene (tk-neo) linked to the hLAR became inducible when stably transfected into MEL cells (12, 13). A condensed version of the hLAR containing all four hypersensitive sites in a 2.5-kb fragment (μLAR) (14) also conferred full inducibility on a linked human β-globin gene stably transfected into MEL cells; the DNase I-hypersensitive sites of the LAR were reestablished in these cell lines. Subdivision of the hLAR sites indicated that pairs of sites conferred some inducibility on linked globin genes in MEL cells, but at least three of the four sites are required for full inducibility (15). Finally, Driscoll *et al.* (16) have described a patient heterozygous for γδβ-thalassemia who carries a deletion of LAR sites II–IV and additional upstream sequences; this deletion inactivates all of the globin genes on the affected chromosome, suggesting that the LAR is important for human globin gene function *in vivo*.

In this report, we describe the molecular cloning of murine DNA sequences homologous to hLAR site II.† These sequences are linked to the mouse β-globin gene cluster in the same basic arrangement as the human β-globin gene cluster. The mouse LAR (mLAR) and hLAR share primary DNA sequences and several enhancer-type functions. These data suggest that the conserved regions may be important for function of the LAR.

MATERIALS AND METHODS

Library Screening. A Stratagene λFix murine genomic library was plated, and ≈600,000 plaques were transferred to nitrocellulose filters. The filters were prehybridized for 2 hr at 42°C in 10% dextran sulfate/4× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7)/7 mM Tris·HCl, pH 7.6/0.8× Denhardt's solution (1× Denhardt's = 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin)/40% formamide/denatured, sonicated salmon sperm DNA at 25 μg/ml. The filters were hybridized in the above buffer with 5×10^6 cpm of random primer ^{32}P-labeled hLAR site II probe (a *Hin*dIII–*Bgl* II fragment extending from −11,054 to −10,322 with respect to the human ε-globin gene cap site) per ml. The filters were washed twice at room temperature in 2× SSC/0.1% SDS and twice at 55°C in 0.1× SSC before exposure to x-ray film overnight.

K562 Cell Transfections. For transient transfections, K562 cells were electroporated with a capacitance of 800 μF at 275 V essentially as described (17). Total cellular RNA was harvested from the cells 20–24 hr after electroporation, hybridized, and analyzed as described (17).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: LAR, locus-activating region; h, human; m, mouse; SV40, simian virus 40; nt, nucleotide(s).

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M37325).

Electroporation conditions for the generation of stable transformants were 600 μF and 200 V. For the colony assay, cells were incubated without selection in complete medium [Iscove's modified Eagle's medium with 5% fetal bovine serum and 5% controlled process serum replacement type 4 (Sigma)] for 24 hr, diluted, and added to ≥10 volumes of prewarmed complete medium containing 0.3% agarose and G418 (GIBCO) at 1 mg/ml. All experiments were plated in duplicate. Visible colonies were quantitated 2 weeks after plating.

RESULTS AND DISCUSSION

Globin LAR Sequences Are Conserved in the Mouse Genome and Are Linked to the Mouse β-Globin Gene Cluster. The transgenic mouse and MEL cell experiments demonstrated that murine erythroid cells have the factors required for appropriate function of the hLAR. This suggested to us that mice may have similar LAR elements and that conservation between hLAR and mLAR counterparts would identify regions important for LAR function. We therefore performed low-stringency Southern hybridization with blots containing restricted mouse genomic DNA, using a probe derived from hLAR site II (see *Materials and Methods*). We detected specific murine fragments (e.g., 7-kb *Bam*HI, 1.1-kb *Eco*RI, 1.1-kb *Hin*dIII, and 10-kb *Pst* I), indicating that discrete regions in the mouse genome are related to hLAR site II (data not shown). We next screened a mouse genomic library using identical hybridization conditions and isolated a clone (λmLAR) containing a 13-kb insert (Fig. 1). Southern analysis of restricted λmLAR DNA probed with hLAR site II revealed that λmLAR had most of the restriction sites predicted by the initial mouse genomic digests. Mapping of λmLAR indicated that all of the homology with hLAR site II was contained within a 1.1-kb *Eco*RI fragment (mLAR1.1, Fig. 1). Additionally, a probe derived from mLAR1.1 cross-hybridized with the same restriction fragments in human genomic DNA detected by the hLAR site II probe. Sequence analysis of mLAR5 (see Fig. 1) established linkage with the mouse β-globin cluster; sequence that we obtained corresponded exactly with published sequences from positions −9340 to −9265 with respect to the mouse εy-globin gene cap site (18). The total overlap between λmLAR and the published sequence upstream from the mouse β-globin cluster is 1.7 kb (see Fig. 1).

```
TTTCCTTACT TCTTTTTTTT CTTCCTTCAT TTTGAGTCAG GATTTTACTA TATAACTATG   -10670
CTATCATGGA ACATACTATG TAGATCAGGA TTGACTGGTA CTCACAGAGA TCTGCCTGTC   -10610
CCTGCCTCGT GAGGTCTTAG ACTAAAGATG TGACCTCCTT CACTCTATAT CCAACTAAAA   -10550
CGCTGACCAC TGATGGCTTA GTTTAACTAG Aaatgggctt gaattcattg cagactttca   -10490
ctggtaccct gtttccttat ctgacctgct ttaactgggt aagcttatga aaagtcttgt   -10430
gtagaagaga aagggataac agcctgtgCT AAATGAGGAA GCTGCCTTGC CTGTTCCTGC   -10370
TCAGTGGGGT TTCTGGCTAT ACTACATCAA CTGAGTCAGT GCATCTTGAC TAGTTCCCAC   -10310
ACCTTTCTCT GAACAGAGAG AGTAAAGGGC TCAATAAGAA AAATACAGTT TATGGTCTGT   -10250
ACGTTGTATT ACACGTCATC CCTTACTTTC TAAAAGGCAT CTTCACTGAG AAAGACATGG   -10190
ATTTCTAACC TTACAATCTC TACAAAGTTC AGAATACATA ATAATCTTGA ATTATGATTA   -10130
ACTGTTAGTT TTGACCAGGT CTTCTGGCAG ACAGGTCACA TGTGTTAGTA TCACTTATTC   -10070
CTCAAGTGTT GATGTTAGTG TCAGCATATT ACCGATGTTC CACAAACATT TCCTGAATGA   -10010
CTGTTAAACT TCCTACACAT TAACGAGCCT CTGCATTTTT TTCCAGCTTC CATCTATGAT    -9950
TTAAGTAAAC TCTAGTTTTC CACTTCTTCA TATTCTCTCT CTAGATCTCA ATTATTGCAG    -9890
TACCACTGTC CAAGGGCAGA GGAGGTTAGC TGGGCCCAGG CGGAGTCAAT TCTCTACTCC    -9830
CCACCCTGTG GGTGTGTTCA GCCTTGTGAG CCAGCATCAG GCTTGAGCAC AGCAGTGCTG    -9770
AGTCATGCTG AGTCATGCTG AGGCTTAGGG TGTGTGGCCA GATGTTTTCA GCTGTGAGTG    -9710
ATCAGTGCTA TCTGGGTCTC TAGGAGGAAG TCCACAGGGA AGGTGAAAAG AAAATAAGTT    -9650
TGCTCCCTGA AGAAAACATT ACTTACACCA GCATTACAAT GAAAAGGGGA CCCTGCCTTG    -9590
CTGTGTGACA TAACCTAGAA TATTTTATTT TCTAGTTAAA AATTACTTCT CATGATCTTT    -9530
ATTAGAAGTT ATAACAAATC ATTTTTTAGA AATGTAAAGT TTATTTAGAC TCATACCTAG    -9470
AAGCAACAGA ATCACACATT TTAGTGTACA CACACACACA CACACGCACA CACACACACA    -9410
TGACTTTTGT TCATAAACTA TGCAATTGTT AATGAAATGC TATTTGGaat gggcttgaat    -9350
tcattgcaga cattcactgg tacctgttt ccttatctga cctgctttaa ctgggtaagc    -9290
ttatgaaaag tctggtgtag aagagaaagg ggtaataacc tgtgTTGTTA GAGTTAGGGT    -9230
CATTAGATCT                                                          -9220
```

FIG. 2. Sequence of mLAR site II from −10,729 to −9220 with respect to the mouse εy-globin gene cap site. The entire sequence was determined on both strands using standard techniques (except for the first 100 bp, which was determined on one strand only) (17). The directly repeated sequences are shown in lowercase letters. The locations of the tandem AP-1 sites and poly(pyrimidine-purine) tract are underlined. The *Eco*RI sites that define the mLAR1.1 fragment are located at positions −10,509 and −9353.

hLAR and mLAR Site II Regions Contain Motifs with Striking Homology. The sequence of mLAR site II (Fig. 2) contains a region of >70% identity with hLAR site II extending over ≈800 base pairs (bp) (see Fig. 3). Multiple regions with >80% identity are scattered throughout this region (see Figs. 2 and 3). A core with 95% identity over 60 bp is centered over two tandem transcription factor AP-1 consensus elements (19, 20) located at position −10,879 relative to the human ε-globin gene cap site or position −9771 relative to the mouse εy-globin gene cap site. Several investigators have recently demonstrated that this region is important for the enhancer-like function of the hLAR (21–23). The sequence TGCTGAG is repeated three times in tandem over the AP-1 elements; this motif is also preserved in human sequences. mLAR site II also contains scattered regions of homology with the simian virus 40 (SV40) enhancer. For

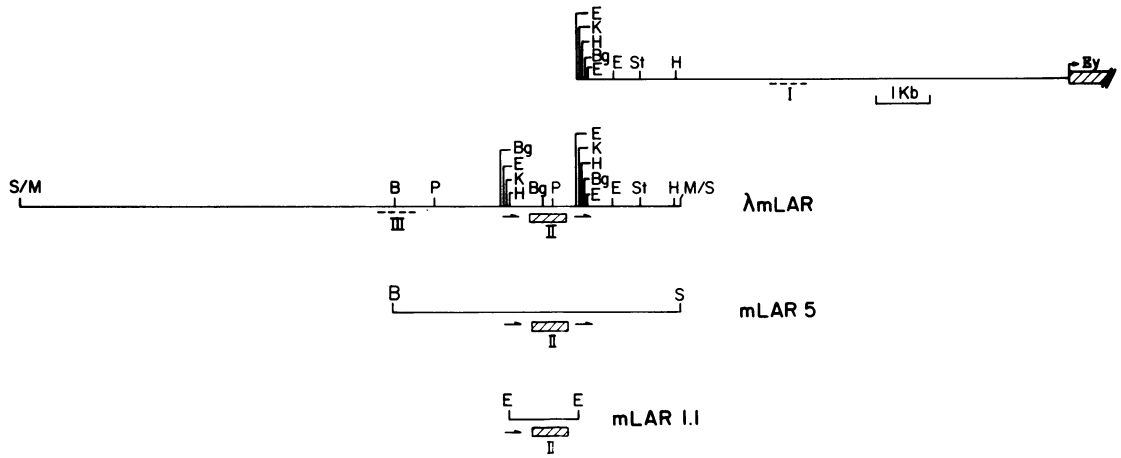

FIG. 1. Maps of the region upstream from the mouse β-like globin gene cluster. The top diagram represents previously published sequence (18) and a partial restriction map of the region 5′ to the mouse εy-globin gene. A map of the mouse λmLAR clone is shown on the next line; restriction sites that overlap with the previously published sequence are aligned. The restriction map shows only selected sites and is incomplete upstream from the *Bam*HI site that forms the 5′ end of mLAR5. The region containing strong homology with the human β-globin LAR site II is indicated by a hatched box. The mouse LAR directly repeated sequences are designated by arrows. Regions with homology to human β-globin LAR sites I and III are shown. Subfragments of λmLAR used in the functional studies (mLAR5 and mLAR1.1) are designated. Note that mLAR1.1 contains the entire region of strong homology with hLAR site II. E, *Eco*RI; S, *Sal* I; M, *Mbo* II; P, *Pvu* II; K, *Kpn* I; B, *Bam*HI; Bg, *Bgl* II; H, *Hin*dIII; St, *Sst* I.

example, there is an SV40 enhancer core-like element (24) in mLAR site II from position −9745 to −9731 and in hLAR site II from position −10,850 to −10,839. Another conserved feature in this region is a poly(pyrimidine-purine) stretch consisting of T and A residues in the human (position −10,656 to −10,607) or C and A residues in the mouse (position −9442 to −9410).

In the mouse sequence, we discovered direct repeats of 117 bp positioned at −10,518 and −9362 (see Fig. 2). The two copies of the direct repeat are indicated in lowercase letters in Fig. 2 and differ from each other at five positions. One copy of this direct repeat is located in the hLAR site II at position −11,924 (see Fig. 3); this sequence is 80% identical with the mouse direct repeat over a stretch of 85 bp. However, the 3′ copy of the direct repeat is not present in hLAR site II. A search of GenBank data base release no. 61 (October 1989) revealed no other sequences with significant homology to the mouse direct repeats.

The LAR site II also contains several consensus binding sites for the erythroid/megakaryocytic transcription factor GF-1 (NFE-1 or Eryf-1) (25, 26). One site, conserved in location between mouse and human sequences, is found on the "noncoding" strand of mLAR at position −9698 and the noncoding strand of hLAR at position −10,802. The mLAR direct repeats contain GF-1 sites on the noncoding strand at positions −10,469 and −9312; however, the human copy of the direct repeat does not contain the GF-1 site. Additional GF-1 consensus sites are found in the hLAR (but not the mLAR) at positions −10,543 (noncoding strand) and −11,699 (coding strand). The significance of all of these consensus sites is unknown at this time.

LAR Sites I and III Are also Conserved in the Mouse. Conservation of the location of mLAR site II relative to the mouse β-like globin gene cluster suggested that the entire region upstream from the human and mouse β-like globin clusters may be conserved. We therefore performed a homology comparison between the published mouse (18) and human (27) sequences upstream of the εy- and ε-globin genes, respectively, and found that the region from approximately −6240 to −5540 upstream from the εy-globin gene was ≈65% identical with the region between approximately −6440 and −5740 with respect to the human ε gene. The positions and sequences of LAR sites I and II are therefore conserved between mice and humans. Additionally, since we knew that mLAR5 was located at the extreme 3′ end of λmLAR, we suspected that the murine counterparts of hLAR sites III and IV might also be present in λmLAR. We therefore blotted restriction enzyme-cleaved λmLAR DNA and probed the filter with hLAR fragments derived from site III or IV (14). The human site III probe hybridized with DNA fragments located just upstream from site II (data not shown). The human site IV probe hybridized only weakly with λmLAR sequences even further upstream. The proposed locations of the mLAR sites are shown in Fig. 1.

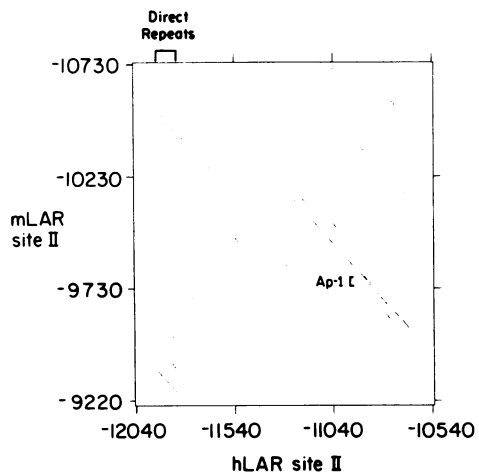

FIG. 3. Dot matrix comparison of mLAR site II (−10,729 to −9220 with respect to the mouse εy-globin gene cap site) and the hLAR site II region (−12,040 to −10,540 with respect to the human ε-globin gene cap site). The stringency of the comparison included a minimum match of at least 11 of 15 bp. The locations of the mouse LAR direct repeats, the human LAR direct "repeat," and the tandem AP-1 sites are indicated.

mLAR Site II Is a Transcriptional Enhancer. We next tested the function of the mLAR by using assays previously developed to study function of the hLAR. To test for the ability of the hLAR and mLAR to act as transcriptional enhancers in a transient expression system, we inserted the human μLAR (hμLAR) fragment or mLAR1.1 (see Fig. 1) into γ-neo, a pUC9-based plasmid containing an ^{A}γ-globin promoter (extending from −299 to +36 with respect to the γ-globin gene cap site) driving the neomycin phosphotransferase gene. K562 cells were electroporated with these plasmids as described (17). Approximately 24 hr after transfection, total cellular RNA was harvested and hybridized with end-labeled probes derived from γ-neo and an endogenous γ-globin gene (17). Samples were then treated with S1 nuclease and analyzed on sequencing gels. Correctly initiated γ-neo mRNA protects a γ-neo probe fragment of 375 nucleotides (nt) from S1 nuclease digestion; correctly spliced endogenous γ-globin exon 2 mRNA protects a γ-globin probe fragment of 209 nt. All points were cotransfected with an internal control plasmid (RSV-neo, which contains the Rous sarcoma virus long terminal repeat promoter driving neo) to control for transfection efficiency. The RSV-neo mRNA protects a γ-neo probe fragment of 325 nt (see Fig. 4 and ref. 17). Fig. 4 shows the results of a representative S1 nuclease protection assay using the indicated γ-neo plasmids. Densitometric analysis of the data presented in this figure revealed that the human μLAR increased correctly initiated γ-neo mRNA levels by ≈3.8-fold and that mLAR 1.1 increased levels by ≈3.3-fold (Fig. 4, compare lanes 2 and 3 with lanes 5 and 6). The effect of the LAR fragments on γ-neo mRNA levels was position and orientation independent.

mLAR and hLAR Sequences Increase the Number of Productive Genomic Integration Events in K562 Cells. In experiments with populations of K562 cells stably transfected with γ-neo versus γ-neo/hμLAR (plasmid designations are explained in the legend to Table 1), we observed that many more G418-resistant colonies were obtained when the hμLAR was linked with γ-neo (see Table 1). However, measurement of correctly initiated γ-neo mRNA from pools of clones containing γ-neo (Fig. 5, lanes 1 and 4) versus γ-neo/hμLAR (Fig. 5, lanes 2 and 3) revealed that the level of correctly initiated γ-neo mRNA is approximately equal in each population. This suggests that a "threshold" level of γ-neo mRNA is required for G418-resistant colonies to form; however, the hμLAR does not further enhance γ-neo mRNA levels over that seen with γ-neo alone. To quantify the LAR effect, we used a simple assay for colony formation in semisolid G418-containing media (28, 29). K562 cells were electroporated with the indicated plasmids and then plated in duplicate, as described in *Materials and Methods*. Visible colonies were quantitated 2 weeks after plating. Table 1 contains the results obtained from a number of such experiments. The γ-neo/hμLAR plasmids yielded ≈70 times as many G418-resistant colonies per 10^6 transfected cells as γ-neo alone; the γ-neo/hLAR site II plasmids yielded approximately the same number of G418-resistant colonies as the entire hμLAR. We tested γ-neo/mLAR5 and γ-neo/mLAR1.1 and found that both mLAR fragments also had a marked effect on the number of G418-resistant colonies obtained. The magnitude of the effect with hLAR and mLAR fragments was compa-

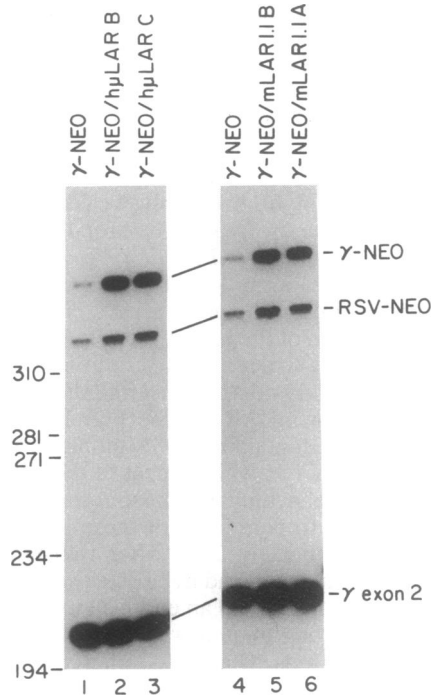

FIG. 4. S1 nuclease protection assay of total RNA prepared from transiently transfected K562 cells. Approximately 7.5 × 10^6 K562 cells in late logarithmic phase growth were transfected with 10 μg of the supercoiled γ-neo plasmid (or the molar equivalent of γ-neo/hμLAR or γ-neo/mLAR1.1) and 2 μg of RSV-neo as described (17). Twenty-four hours after transfection, total cellular RNA was harvested and hybridized with 100,000 cpm (≈10 ng) of ^{32}P-end-labeled probes derived from γ-neo and an endogenous γ-globin gene. After S1 digestion, probe fragments were resolved on sequencing gels and autoradiographed overnight at −70°C. The sizes of correctly initiated γ-globin mRNA, RSV-neo mRNA, and endogenous γ-globin exon 2 mRNA are described in the text and indicated at the right. The positions of simultaneously run molecular markers (in nt) are shown at the left. The γ-neo/hμLAR constructs (lanes 2 and 3) contain the 2.5-kb μLAR fragment (14). The γ-neo/mLAR1.1 plasmid contains the 1.1-kb *Eco*RI fragment derived from λmLAR (Fig. 1). Note that both hLAR and mLAR sequences increase the levels of correctly initiated γ-neo mRNA to a similar extent, regardless of the orientation of the LAR fragment in the plasmid. Orientations of the LAR fragments with respect to the γ-neo transcription unit are indicated by the letters following the plasmid name as follows: A, upstream tandem; B, upstream opposed; C, downstream tandem.

rable. A variety of control DNA fragments of similar size (e.g., human LAR site IV, see Table 1) had no effect in the colony assay. Finally, preliminary experiments have demonstrated that the hμLAR has no effect on colony formation in murine 3T3 cells, suggesting that this effect may be specific for erythroid cells (M. J. Ulrich and T.J.L., unpublished observations).

hLAR and mLAR Sequences Confer Hemin Inducibility on Linked γ-Globin Promoters. We next wished to test the capacity of the hLAR or mLAR sequences to render a linked γ-globin promoter hemin inducible in stably transformed K562 cells (21, 22). We therefore pooled several K562 cell clones stably transformed with either γ-neo, γ-neo/hμLAR, or γ-neo/mLAR5. These pools of clones were induced with 20 μM hemin for 5 days; total cellular RNA derived from these pools was analyzed for the presence of correctly initiated γ-neo mRNA and endogenous γ-globin mRNA (Fig. 5, lanes 7–12). Comparison of γ-neo mRNA levels in uninduced and induced K562 cells reveals that γ-neo itself is not inducible, even though endogenous γ-globin levels increased ≈8-fold with hemin treatment (Fig. 5, compare lanes 1 and 4 with lanes 7 and 10). These data are consistent with that

Table 1. Colony formation in stably transfected K562 cells

Transfected plasmid	Colony number per 10^6 transfected cells (mean ± SD)	*n*	Fold increase relative to γ-neo
None	<1	2	—
γ-neo	20 ± 9	4	—
γ-neo/hμLAR B	1675 ± 601	2	84
γ-neo/hμLAR C	1345 ± 220	2	67
γ-neo/hLAR II A	1230 ± 770	3	65
γ-neo/hLAR II B	1200 ± 510	3	63
γ-neo/hLAR IV B	26 ± 7	2	<2
γ-neo/hLAR IV C	28 ± 10	4	<2
γ-neo/mLAR5 B	220 ± 58	2	12
γ-neo/mLAR5 C	1080 ± 190	2	57
γ-neo/mLAR1.1 B	620 ± 31	2	33
γ-neo/mLAR1.1 C	1000 ± 28	2	52

For each experiment, ≈7.5 × 10^6 K562 cells in late logarithmic phase growth were electroporated with 10 μg of the supercoiled pUC9-based γ-neo plasmid or an equimolar amount of LAR-containing plasmids. The γ-neo/hμLAR plasmids contain the 2.5 kb μLAR fragment described in ref. 14, the γ-neo/hLAR II plasmids contain a 0.73-kb *Hin*dIII–*Nco* I fragment extending from −11,054 to −10,322 with respect to the ε-globin gene cap site, and γ-neo/hLAR IV plasmids contain an *Eco*RI–*Apa* I fragment derived from the μLAR extending from −18,589 to −17,838 (14, 27). The γ-neo/mLAR plasmids contain mLAR fragments defined in Fig. 1. The letters following the plasmid names refer to the orientation of the inserts with respect to the γ-neo transcription unit as defined in Fig. 4. The total number of colonies per 10^6 transfected cells per μg of transfected γ-neo plasmid DNA is similar to that observed in previous studies (28, 29).

previously described by Acuto *et al.* (29). However, clones containing γ-neo/hμLAR demonstrate inducibility of γ-neo mRNA with hemin (Fig. 5, compare lane 7 with lane 8 or 9). Similarly, addition of the mLAR5 fragment to γ-neo confers hemin inducibility on the linked γ-neo transcription unit, as demonstrated in lane 10 versus lane 11 or 12. Densitometric analysis of the data presented in Fig. 5 revealed that the hμLAR increased correctly initiated γ-neo mRNA levels ≈13-fold with hemin induction and that the effect with mLAR5 was ≈11-fold. The effect of mLAR1.1 was equivalent to that of mLAR5 (data not shown). A separate Southern blot analysis revealed that the average copy number of γ-neo in each pool was approximately equivalent (data not shown).

Our functional analyses of the hLAR and mLAR corroborate several previous studies of LAR function. In our experiments, the hLAR and mLAR elements both acted as transcriptional enhancers when transiently transfected into uninduced K562 cells, a result similar to that described by Nienhuis and coworkers (21, 22) as well as Moi and Kan (23). In their studies, the AP-1 region of hLAR site II functioned as a transcriptional enhancer in uninduced K562 cells; the enhancer effect was induced with hemin. The colony assay we describe here was also performed in uninduced K562 cells. Our results suggest that the factors required for the LAR function detected by the colony assay are present in uninduced cells, consistent with the fact that the DNase I-hypersensitive sites of the hLAR are present in uninduced K562 cells (3). However, when we examined the levels of correctly initiated γ-neo mRNA in pools of stably transfected, uninduced K562 cells, the LAR did not increase the level of correctly initiated γ-neo mRNA.

These observations are perhaps best understood in light of previous experiments by Moreau *et al.* (30) and Weintraub (31), who demonstrated that the SV40 enhancer increased the *percentage* of transiently transfected cells that expressed detectable amounts of large tumor antigen; however, the SV40 enhancer did not appear to affect the *level* of SV40

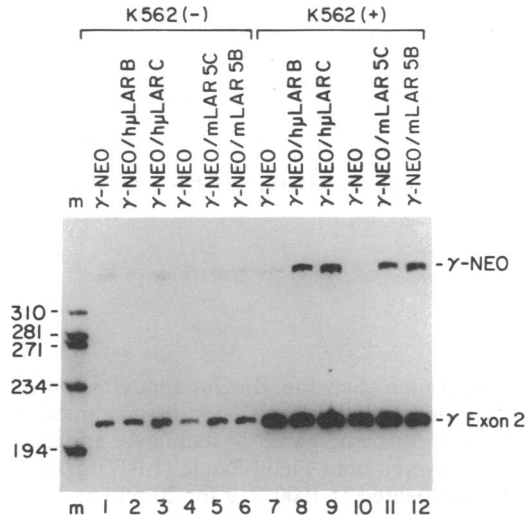

FIG. 5. S1 nuclease protection assay of total RNA derived from pools of stably transformed K562 cells uninduced (lanes 1–6) or induced (lanes 7–12) for 5 days with 20 μM hemin. Ten micrograms of total cellular RNA derived from each indicated pool was hybridized with ≈100,000 cpm (≈10 ng) of the γ-neo and endogenous γ-globin probes. Samples were analyzed as described in the legend to Fig. 4. The positions of probe fragments corresponding to correctly initiated γ-neo mRNA and correctly spliced γ-globin exon 2 mRNA are shown. The sizes of simultaneously run molecular markers (in nt) are shown in the lane marked "m." Note that levels of correctly initiated γ-neo mRNA in uninduced cells are approximately the same regardless of the presence of a LAR. However, treatment with hemin causes increased levels of γ-neo mRNA to accumulate only in pools transfected with hμLAR (lanes 8 and 9) or mLAR5 (lanes 11 and 12) containing plasmids, even though endogenous γ-globin mRNA increases in all pools (lanes 1–6 versus lanes 7–12). Plasmid designations are as described in the legend to Fig. 4.

promoter activity. In our transient transfection experiments, we propose that the LAR elements act to increase the percentage of cells that express the γ-globin promoter at a level that is determined by promoter sequences, and therefore we detect more correctly initiated γ-neo mRNA. In contrast, the colony assay detects individual "productive" integration events at a *single* cell level. A threshold level of γ-globin promoter activity is required for survival of a transfected cell in G418-containing selective medium, since γ-globin promoter activity is required for neomycin phosphotransferase production. The LAR elements greatly increase the number of productive integration events (i.e., events that yield threshold levels of γ-neo mRNA), but the levels of correctly initiated γ-neo mRNA in pools derived from γ-neo versus γ-neo/LAR are the same (see Fig. 5, lanes 1–6). These results suggest that the LAR increases colony numbers, not by increasing γ-globin promoter activity, but by increasing the likelihood that each integration event will yield a threshold level of γ-neo mRNA. We therefore propose that the colony assay detects an activity of the LAR that is related to its "domain opening" function. The very few colonies that form with γ-neo alone represent integration events that permit expression of the γ-globin promoter at the threshold level required for survival. However, these integration sites do not contain all LAR type activities, since they do not render the linked γ-globin promoter hemin inducible. Further experiments will be required to dissect these various functions of the human and mouse LAR elements.

We have cloned and characterized murine sequences that are homologous to the hLAR. The mLAR is linked to the mouse β-like globin gene cluster and shares primary structure, organization, and function with its human counterpart. Our results suggest that functional locus-activating elements are defined by primary DNA sequences that are conserved across species. Conservation of the locations and sequences of LAR sites I and III between human and mouse provides evidence that the organization of primary sequence elements may also contribute to LAR function *in vivo*. Further experimentation will be needed to determine whether the most conserved DNA sequences within these elements are the most important for LAR functions.

Note Added in Proof. Since submission of this report, Talbot *et al.* (32) have also described the importance of the AP-1/NF-E2 region for the function of hLAR site II.

We thank Seth Crosby and Kelli Simberger for oligonucleotides, members of the Ley lab for helpful discussions, and Diana Coleman for preparing this manuscript. Kathleen Maloney provided expert technical assistance. William Forrester and Mark Groudine kindly provided the human μLAR plasmid. This work was supported by National Institutes of Health Grant DK 38682 and a March of Dimes/Colonel Sanders Memorial Fund Basil O'Connor Award.

1. Tuan, D., Solomon, W., Li, Q. & London, I. M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6384–6388.
2. Forrester, W. C., Thompson, C., Elder, J. T. & Groudine, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1359–1363.
3. Forrester, W. C., Takegawa, S., Papayannopoulou, T., Stamatoyannopoulos, G. & Groudine, M. (1987) *Nucleic Acids Res.* **15**, 10159–10177.
4. Grosveld, F., van Assendelft, G. B., Greaves, D. R. & Kollias, G. (1987) *Cell* **51**, 975–985.
5. Ryan, T. M., Behringer, R. R., Martin, N. C., Townes, T. M., Palmiter, R. D. & Brinster, R. L. (1989) *Genes Dev.* **3**, 314–323.
6. Ryan, T. M., Behringer, R. R., Townes, T. M., Palmiter, R. D. & Brinster, R. L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 37–41.
7. Hanscombe, O., Vidal, M., Kaeda, J., Luzzatto, L., Greaves, D. R. & Grosveld, F. (1989) *Genes Dev.* **3**, 1572–1581.
8. Behringer, R. R., Ryan, T. M., Reilly, M. P., Asakura, T., Palmiter, R. D., Brinster, R. L. & Townes, T. M. (1989) *Science* **245**, 971–973.
9. Greaves, D. R., Fraser, P., Vidal, M. A., Hedges, J. J., Ropers, D., Luzzatto, L. & Grosveld, F. (1990) *Nature (London)* **343**, 183–185.
10. Curtin, P. T., Liu, D., Liu, W., Chang, J. C. & Kan, Y. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7082–7086.
11. Tuan, D., Solomon, W. B., London, I. M. & Lee, D. P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2554–2558.
12. van Assendelft, G. B., Hanscombe, O., Grosveld, F. & Greaves, D. R. (1989) *Cell* **56**, 969–977.
13. Talbot, D., Collis, P., Antoniou, M., Vidal, M., Grosveld, F. & Greaves, D. R. (1989) *Nature (London)* **338**, 352–359.
14. Forrester, W. C., Novak, U., Gelinas, R. & Groudine, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5439–5443.
15. Collis, P., Antoniou, M. & Grosveld, F. (1990) *EMBO J.* **9**, 233–240.
16. Driscoll, M. C., Dobkin, C. S. & Alter, B. P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7470–7474.
17. Ulrich, M. J. & Ley, T. J. (1990) *Blood* **75**, 990–999.
18. Sheehee, W. R., Loeb, D. D., Adey, N. B., Burton, F. H., Casavant, N. C., Cole, P., Davies, C. J., McGraw, R. A., Schichman, S. A., Severynse, D. M., Voliva, C. F., Weyter, F. W., Wisely, G. B., Edgell, M. H. & Hutchison, C. A., III (1989) *J. Mol. Biol.* **205**, 41–62.
19. Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P. & Karin, M. (1987) *Cell* **49**, 729–739.
20. Lee, W., Mitchell, P. & Tjian, R. (1987) *Cell* **49**, 741–752.
21. Ney, P. A., Sorrentino, B. P., McDonagh, K. T. & Nienhuis, A. W. (1990) *Genes Dev.* **4**, 993–1006.
22. Sorrentino, B., Ney, P., Bodine, D. & Nienhuis, A. W. (1990) *Nucleic Acids Res.* **18**, 2721–2731.
23. Moi, P. & Kan, Y. W. (1989) *Blood* **74**, 7a (abstr.).
24. Banerji, J., Olson, L. & Schaffner, W. (1983) *Cell* **33**, 729–740.
25. Tsai, S.-F., Martin, D. I. K., Zon, L. I., D'Andrea, A. D., Wong, G. G. & Orkin, S. H. (1989) *Nature (London)* **339**, 446–451.
26. Trainor, C. D., Evans, T., Felsenfeld, G. & Boguski, M. S. (1990) *Nature (London)* **343**, 92–96.
27. Li, Q., Powers, P. A. & Smithies, O. (1985) *J. Biol. Chem.* **260**, 14901–14910.
28. Rutherford, T. & Nienhuis, A. W. (1987) *Mol. Cell. Biol.* **7**, 398–402.
29. Acuto, S., Donovan-Peluso, M., Giambona, N. & Bank, A. (1987) *Biochem. Biophys. Res. Commun.* **143**, 1099–1106.
30. Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M. P. & Chambon, P. (1981) *Nucleic Acids Res.* **9**, 6047–6068.
31. Weintraub, H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5819–5823.
32. Talbot, D., Philipsen, S., Fraser, P. & Grosveld, F. (1990) *EMBO J.* **9**, 2169–2178.