Nucleic Acids Research, 1996, Vol. 24, No. 13 **2511–2518**

Genetic regulation of δ-aminolevulinate dehydratase during erythropoiesis

Terry Rogers Bishop[1,2,*], Miles W. Miller[3,+], Jennifer Beall[1], Leonard I. Zon[4] and Peter Dierks[3,§]

[1]Department of Pediatric Hematology, [2]Center for Medical Genetics, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA, [3]Department of Biology, University of Virginia, Charlottesville, VA, USA and [4]Howard Hughes Medical Institute and Division of Hematology, Children's Hospital, Harvard Medical School, Boston, MA 02115, USA

Received April 3, 1996; Revised and Accepted May 17, 1996

ABSTRACT

In an effort to understand how the heme biosynthetic pathway is uniquely regulated in erythroid cells, we examined the structure of the gene encoding murine δ-aminolevulinate dehydratase (ALAD; EC4.2.1.24), which is the second enzyme of the pathway. The gene contains two first exons, named 1A and 1B, which are alternatively spliced to exon 2, where the coding region begins. Each first exon has its own promoter. The promoter driving exon 1A expression is TATA-less and contains many GC boxes. In contrast, the exon 1B promoter bears regulatory sequences similar to those found for β-globin and other erythroid-specific genes. Tissue distribution studies reveal that ALAD mRNA containing exon 1A is ubiquitous, whereas mRNA containing exon 1B is found only in erythroid tissues. This finding, together with our further observation that GATA-1 mRNA levels increase 3-fold during maturation of murine erythroid progenitor cells, may help explain simultaneous 3-fold increases in exon 1B expression. The unexpected result that exon 1A expression also increases 3-fold during CFU-E maturation may be attributable to the action of NF-E2, since there is a potential binding site in a position analogous to the NF-E2 site in the locus control region of the β-globin gene cluster.

INTRODUCTION

The level of heme synthesized in erythroid cells is orders of magnitude greater than that in other cells (1) and, under normal conditions, is stoichiometrically matched with the levels of globin chain synthesis. It is becoming apparent that the genetic mechanisms responsible for this considerable quantitative difference are often distinctive for each of the eight heme biosynthetic enzymes. For example, the polypeptide chain of the first enzyme in the heme pathway, δ-aminolevulinate synthase (ALAS) is encoded by two divergent genes in chicken (2) and in humans, one ALAS gene, ALAS-N, is located on an autosome (3) and the other, ALAS-E, is X-linked (4). The autosomal gene is expressed in all tissues and has regulatory protein binding sites similar to those for other housekeeping genes (5) while the X-linked one is expressed only in erythroid cells and has some of the same regulatory sites found for other erythroid-specific genes; namely, GATA-1 and NF-E2 (6,7). In murine erythroleukemia (MEL) cells, it has been shown that there is an inverse relationship between ALAS-N and ALAS-E expression: ALAS-E mRNA increases and ALAS-N mRNA decreases upon DMSO-induction (8).

In contrast, all mRNAs for the third enzyme in the pathway, porphobilinogen deaminase (PBG-D), arise from the same locus (9) but there is differential splicing according to tissue type, leading to housekeeping and erythroid-specific isozymes. A typical TATA-less promoter is found upstream of the housekeeping variant (10) whereas GATA-1 and NF-E2 sites are found upstream of the erythroid-specific version (11,12). Amongst expression patterns for other heme pathway enzymes thus far examined, those for uroporphyrinogen decarboxylase (13), coprophyrinogen oxidase (14), protoporphyrinogen oxidase (15) and ferrochelatase (16), the fifth, sixth, seventh and eighth enzymes, respectively, appear to have messengers and polypeptide structures alike in all tissues. Although it is unclear what mechanisms are responsible for augmented expression of uroporphyrinogen decarboxylase (17) and coprophyrinogen oxidase (14) in erythroid cells, mechanisms for differential tissue expression of ferrochelatase include a multipurpose promoter region, which includes not only binding sites for housekeeping transcriptional factors, but also ones for GATA-1 and NF-E2 (18) and a downstream erythroid repressor (19).

In this report, we show how the regulation of the second enzyme in the heme pathway, δ-aminolevulinate dehydratase (ALAD), illustrates yet another way in which differential expression in erythroid and non-erythroid cells can be accomplished. We find that expression of ALAD exons is alike in all tissues except for two untranslated first exons, 1A and 1B. These are differentially spliced to the second exon where the translation start signal is located. Consequently, while ALAD enzyme is identical in all tissues, ALAD mRNA occurs in housekeeping (1A) and erythroid-specific (1B) forms. In both man and mouse, the promoter region upstream of exon 1B contains GATA-1 sites, a

* To whom correspondence should be addressed at: Johns Hopkins Univerrsity, Ross 1125, Rutland Avenue, Baltimore, MD 21205, USA

Present Addresses: +Department of Pathology, University of Washington, Seattle, WA 98195, USA and §American Cyanamid, Princeton, NJ 08543, USA

result which may help explain observed increases in ALAD-1B mRNA during erythropoiesis. In contrast, while the promoter region immediately 5′ of exon 1A resembles that of other housekeeping genes, the region 2.3 kb upstream of exon 1A bears a site which may be bound by the erythroid regulatory protein, NF-E2. This site, if it functions in a manner analogous to the enhancer NF-E2 binding site found 50 kb upstream of the human β-globin gene (20), could help account for our finding that ALAD-1A mRNA also increases during erythropoiesis.

MATERIALS AND METHODS

Culture and induction of cell lines

MEL cell lines, M18b (21) and 270-2 (22) were cultured with 10% fetal bovine serum (Hyclone) in Minimum Essential Medium-α (Life Technologies) and induced to differentiate with 1.5 or 2% DMSO, respectively.

Preparation and culture of murine CFU-E

CFU-E were enriched by thiamphenicol treatment of BALB/cByJ mice (Jackson Laboratory) and purified by centrifugal elutriation and Percoll density gradients as previously described (23–25). CFU-E were cultured at 37°C under 5% CO_2, in media containing 30% fetal bovine serum and, as indicated recombinant human erythropoietin (gift of Genetics Institute). Morphologically, CFU-E purity at the outset was 85–90% in Wright-stained cytospins (Shandon) and 80% of these formed 32-cell erythroid colonies after 40 h (26).

Recombinant DNA techniques

We used an SM/J mouse genomic library (gift of Steven Weaver, University of Chicago) constructed by partial *Mbo*I digestion of genomic DNA, followed by size-selection and ligation into the *Bam*HI site of λL47.1. Phage from this library were grown on *E.coli* LE392 and screened according to Maniatis *et al.* (27), using both ALAD3 and ALAD7 rat cDNA probes (28). Hybridizing plaques were purified by three successive rounds of plating and rescreening. Recombinant λ DNA was purified by the plate-lysate method of Maniatis *et al.* (27). Gel-purified restriction fragments of λmALAD-1 and λmALAD-2 were subcloned into pIBI30, pIBI31 (International Biotechnologies), pEMBL18 (gift from S. Lazarowitz, Carnegie Institution, Baltimore, MD) or pBluescriptIIKS$^-$ (Stratagene). Standard procedures were used for all subsequent nucleic acid manipulations (27). Both strands were used to determine nucleotide sequences using Sequenase (US Biochemical) chain termination reaction as suggested by the manufacturer.

RNA sources and nuclease protection assays

RNA was harvested either by the method of Ullrich *et al.* (29) or Chomczynski and Sacchi (30) using one additional phenol/$CHCl_3$ (1:1) extraction. RNA from yolk sacs was derived from 9.5 day post coitum F2 embryos of BALB/c × C57Bl/6J matings. Fetal liver RNA from 14.5 day embryos was provided by Dave Bodine (NIH). S1 nuclease assays (31,32) were performed following overnight hybridization of 50 µg total RNA and 30 fmol ^{32}P end-labeled DNA probe. Non-hybridized probe and single-stranded carrier DNA were digested at 30°C with 300 U/ml S1 nuclease for 90 min. *Sph*I–*Sty*I (690 bp) and *Sca*I–*Bam*HI (420 bp) genomic ALAD fragments (Fig. 1B) were used to protect exon 1A- and exon 1B-containing transcripts, respectively.

For RNase protection assays (33,34), an ALAD exon 1A-specific probe was prepared by subcloning a 160 bp *Rsa*I–*Ava*II fragment from λmALAD-2 (delimited by vertical arrows in Fig. 3) into pIBI31. The orientation of the cloned fragment was determined by nucleotide sequencing. An exon 1B-specific probe was constructed by subcloning a 138 bp *Xba*I–*Pst*I fragment from an A-PCR clone into pIBI31. In later studies, mild Exonuclease III digestion was used to remove the poly(dG) tail from the 5′-end of the exon 1B probe. Removal was verified by nucleotide sequence analysis. Plasmid DNA, purified by CsCl gradient centrifugation and linearized with *Eco*RI, was used to prepare radiolabeled antisense riboprobes with T7 RNA polymerase (Life Technologies).

Each RNase protection assay utilized 5 or 20 µg of test RNA and enough phenol-extracted, ethanol-precipitated *E.coli* tRNA to make the total RNA mass equal to 30 µg, plus 50 000 d.p.m. of radiolabeled antisense riboprobe. RNaseA and RNaseT1 both overdigested the RNA even at the lowest concentrations tested (data not shown), whereas 15 U/ml RNaseT2 left protected fragments intact. We suppose that the successful use of T2 reflects the elevated GC content of ALAD exon riboprobes and a cleavage specificity for A residues.

Anchor-polymerase chain reactions (A-PCR)

The method of Loh *et al.* (35) was used with the following modifications. Poly(A)$^+$ RNA from mouse liver or DMSO-induced MEL cells was used as template for AMV reverse transcriptase (Life Sciences) with oligo(dT) as a primer. After alkali treatment, poly(dG) tails were added to the cDNA using terminal deoxynucleotidyl-transferase (Life Technologies). Tailed cDNA (8 ng) was used in A-PCR with 1 µM ALAD oligonucleotide 25 (oligo 25) and 1 µM of a 1:9 mixture of oligo 2 and oligo 3 anchors. The oligo 25 sequence, spanning residues 368 to 387 in the translated murine ALAD sequence (36), was ATGGGAGGTGTAGGGGCACA; anchor oligo 3 (35) was TGCGGCCGCGGATCCGAATTC; and anchor oligo 2 was the same as oligo 3 plus a poly(C)$_{14}$ tail. *Taq* polymerase amplification (Perkin-Elmer Cetus) was carried out for 25 cycles: denaturing at 94°C for 1 min, annealing at 55°C for 1.5 min and elongating at 72°C for 2.5 min. Products were electrophoretically fractionated on a low-melt agarose gel (FMC). Melted gel slices were subjected to a second round of amplification using oligo 3 and ALAD oligo 24 primers. The oligo 24 sequence, spanning murine ALAD 319–338, was GGGAAGGTCTTCCTCAGCAG. A-PCR products were ligated into *Bam*HI- and *Eco*RV-digested pBluescriptIIKS (Stratagene) and plasmid DNA from clones of transformed *E.coli* HB101 cells analyzed by dideoxynucleotide sequencing.

Northern blot analysis of ALAD mRNA

Poly(A)$^+$ RNA from CFU-E cultured for various times was subjected to electrophoresis in 1% agarose, 2.2 M formaldehyde gels (27); transferred to GeneScreen Plus (DuPont/NEN); and hybridized with ^{32}P-dCTP-oligolabeled probes (Pharmacia) prepared from mouse ALAD subcloned fragments. Blots were melted and rehybridized to radioactive β-actin cDNA (37) and autoradiograph band intensities were quantitated using a PhosphorImager and ImageQuant software (Molecular Dynamics, Sunnyvale, CA 94086).

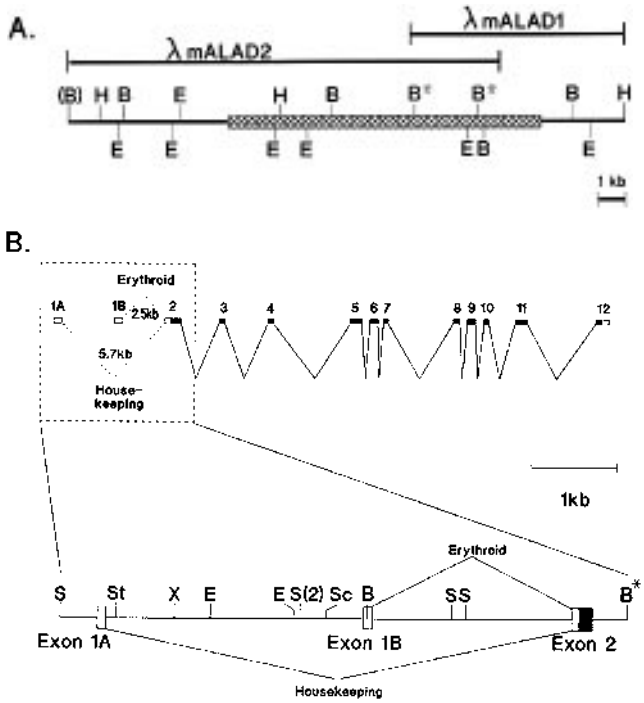

Figure 1. Structure of the murine ALAD gene. **(A)** Restriction map of λmALAD1 and λmALAD2, diagrammed in the 5′ to 3′ direction; B, *Bam*HI; H, *Hin*dIII; and E, *Eco*RI. The *Bam*HI site in parentheses was generated during the cloning process and does not exist in the genome. The hatched box encompasses the 12 kb ALAD gene. Asterisks delimit the 3.1 kb *Bam*HI fragment from λmALAD1 which was used to isolate λmALAD2 with the 5′-most asterisk being located just downstream of exon 2. **(B)** Exon/intron structure of murine ALAD: ■, translated; □, untranslated. Exons 2 through 12 are shown to scale in the upper line and the region on the left is expanded to scale in the lower line; S, *Sph*I; Sc, *Sca*I; St, *Sty*I; X, *Xho*I.

Isolation of the human ALAD gene

Plaque lifts (27) were prepared from a human chromosome 9-specific library (ATCC #57781) and probed with a 350 bp *Eco*RI–*Bam*HI fragment from the mouse ALAD exon 1B promoter (delimited by downward arrows in Fig. 4). The *Eco*RI site was generated by PCR amplification with a primer that mismatched the genomic sequence by one base. Plaques were picked and rescreened until a purified plaque was isolated. Southern blot analysis revealed a 1.2 kb *Sst*I–*Xba*I fragment that contained the human ALAD exon 1B and its promoter and this fragment was subjected to nucleotide sequence analysis (38).

RESULTS

Isolation and structure of the ALAD gene from the SM/J mouse

Approximately 900 000 recombinant plaques were screened using radiolabeled rat cDNA inserts from clones ALAD3 (39) and ALAD7 (28). Three positively hybridizing plaques were isolated. All three recombinant DNAs had identical restriction enzyme maps and were named λmALAD1 (Fig. 1A). A 3.1 kb *Bam*HI fragment from the 5′-end of λmALAD1 (marked by asterisks in Fig. 1A) was used to rescreen an additional 820 000 plaques from the same library. One new clone, out of six acquired, contained a 17 kb insert extending ~11 kb upstream from the 5′-end of λmALAD1 and was named λmALAD2 (Fig. 1A).

cDNA hybridization, subcloning and nucleotide sequence analysis of λmALAD1 and λmALAD2 were used to identify exons 2 through 12 (Fig. 1B). Exon lengths ranged in size from 37 to 187 bp and all splice sites matched the AG/GT rule (40,41). The coding sequence in exons 2 through 12 (data not shown) corresponds exactly to the sequence of mouse reticulocyte ALAD cDNA (36). Analysis of genomic sequence further upstream from the end of the cDNA homology revealed that exon 2 was flanked by a splice acceptor site (Fig. 2, junction of non-shaded and shaded sequences) and not by any obvious transcriptional promoter elements. Such findings suggested the existence of at least one more upstream non-coding exon, which echoed our earlier discovery of a rat ALAD processed pseudogene whose sequence included 160 nondescript nucleotides upstream of what would have been the start of the coding region (42). We therefore supposed that reported ALAD cDNAs from mice (36), rats (28) and humans (43), are in fact truncated copies of the mRNA.

To find the missing upstream exon(s) and identify potential transcriptional regulatory sites, A-PCRs (34) were used to amplify the 5′-end of ALAD mRNA. Among 16 cloned amplificants, nine came from mouse liver and seven from DMSO-induced MEL cells. Surprisingly, two types of A-PCR products were found. In 15 clones, exon 1A was spliced to exon 2. In the remaining clone, derived from MEL cell mRNA, exon 1B was spliced to exon 2 (Fig. 1B). Sequences of exons 1A and 1B, parts of associated introns and their upstream promoter regions are seen in Figures 3 and 4, respectively. While the length of the A-PCR products was variable, none of them contained any material other than exon 1A or 1B at the 5′ terminus. Thirteen of 15 ALAD exon 1A-containing A-PCR products began at position +1 in Figure 3, the others at –21 and –28. All sixteen A-PCR products contained exon 2 sequences shown in the shaded region of Figure 2.

Like typical housekeeping gene promoters (44) the ALAD exon 1A promoter (Fig. 3) lacks a TATA box, contains many GC boxes and, most significantly, lacks the GATA-1, CACCC and GGTGG regulatory sites found in the β-globin locus (45–47). In contrast, such sites not only exist in ALAD exon 1B promoters from both mouse and rat (Fig. 4), but are also spatially arranged the same as they are in β-globin promoters.

ALAD transcriptional start sites and expression patterns in various tissues

As depicted in Figure 5A, a genomic fragment containing ALAD exon 1A protected RNA from brain, liver and spleen from S1-nuclease digestion, whereas an exon 1B-containing genomic fragment protected RNA only in spleen samples. Protected fragment sizes ranged from 37 to 100 bases in the case of exon 1A and from 56 to 87 bases in the case of exon 1B. The prominent smallest bands correspond to positions +1 in Figures 3 and 4.

Similar tissue distribution and sizing results were seen when antisense riboprobes were used to protect RNA from ribonuclease T2 digestion. In Figure 5B, an exon 1A genomic fragment (Fig. 3) was used as a template for riboprobe synthesis and it protected a 37 bp fragment when hybridized to RNA from liver, spleen from anemic animals, lung, kidney and uninduced MEL cells. In contrast, just as with S1-nuclease experiments, radioactive cRNA from exon 1B protected a 112 bp fragment only in spleen RNA

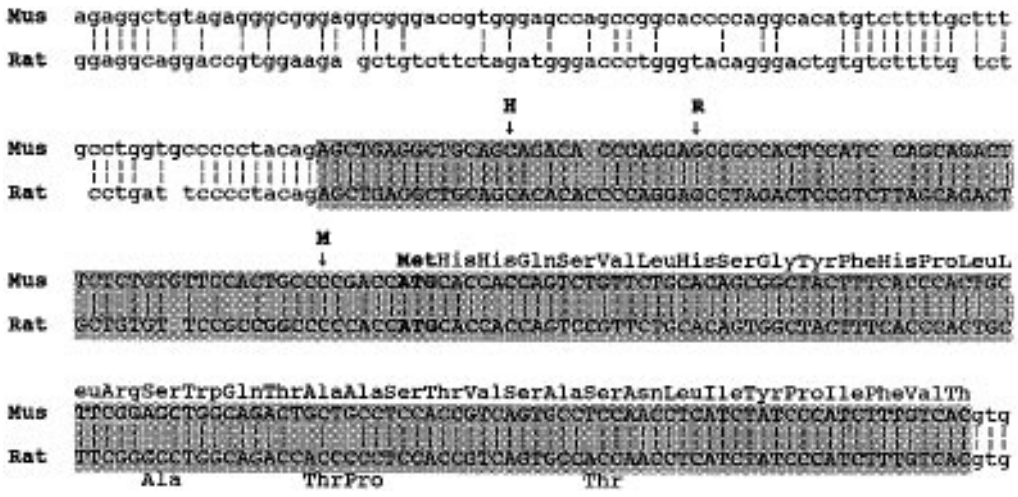

Figure 2. Genomic nucleotide sequences of mouse and rat ALAD exon 2 plus surrounding introns. Shaded, uppercase letters correspond to exon 2 sequences including the ATG translation start site (bold-face) of murine ALAD. Lowercase letters indicate introns. Reported cDNA clones truncate prematurely at down arrows; H, human (43); R, rat (28), and M, mouse (36). The mouse sequence may be accessed through GenBank 191863.

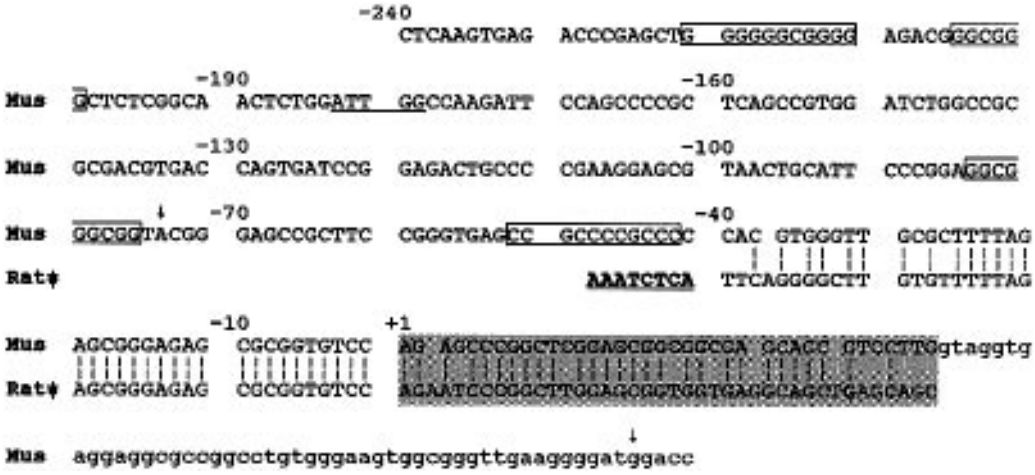

Figure 3. Genomic nucleotide sequences of mouse and rat ALAD exon 1A, upstream proximal promoter regions and downstream intron 1A. Rat ψ refers to the sequence of the rat ALAD processed pseudogene, ψALAD (42). Shaded region denotes the minimal exon 1A, corresponding to the predominant murine start site observed by nuclease protection assays (Fig. 5A and B). Bold underlined letters indicate the direct repeat located at the 5′-end of rat ψALAD (42); lowercase letters, murine intron 1A; boxes, GC motifs; and underline, CCAAT motif, read 5′ to 3′ on the antisense strand. Vertical arrows mark the boundaries of the exon 1A riboprobe used in Figure 5B. The mouse sequence may be accessed through GenBank 191861.

(Fig. 5C), corresponding to transcripts initiating at +1 in Figure 4. The smaller, 82 bp RNase T2-protected fragment arises either from an alternative transcription initiation site or perhaps, more likely, from an alternative splice donor site at position +82, TG/GTAAtT, where the lowercase t indicates a deviation from the consensus splice site. It is not expected that these alternative forms would affect the coding capacity of exon 1B-containing transcripts.

Changes in ALAD erythroid-specific transcript abundance during fetal and adult erythropoiesis

In fetal life (Fig. 6A), the abundance of exon 1B transcripts correlated with the extent of erythropoiesis. They were barely evident in 9.5 d yolk sacs, where erythrocytes are confined to the blood islands, but readily perceptible in 14.5 d fetal liver, where definitive erythropoiesis is abundant. In adults, exon 1B transcripts are scarce in freshly harvested CFU-E (0 h), but then increase to reach a maximum between 21 and 28 h of culture (Fig. 6B), when heme biosynthesis also approaches an apex (26). No significant difference could be observed on the level of exon 1B transcripts in CFU-E grown with or without erythropoietin.

Comparison of human, mouse and rat ALAD erythroid-specific promoter region

A 350 bp fragment containing the mouse erythroid promoter and the 5′ half of exon 1B was used as a probe to isolate the human ALAD clone. Using computer-assisted nucleotide sequence alignments, it was found that only 250 bp of the mouse fragment

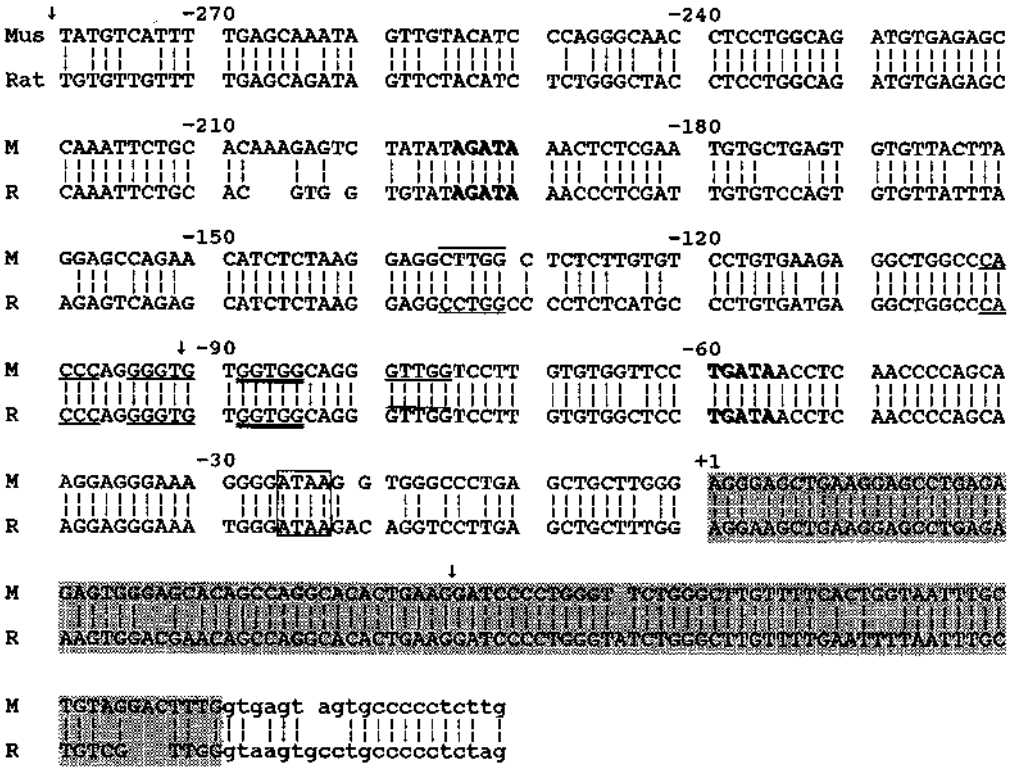

Figure 4. Genomic nucleotide sequences of mouse and rat ALAD exon 1B, upstream proximal promoter regions and downstream intron 1B. The shaded region denotes exon 1B, corresponding to the most prominent fragment detected by RNase T2 and S1 nuclease protection (Fig. 5A and C). Lowercase letters indicate intron; bold-face, GATA-1 sites; overlines, possible CCAAT box read 5′ to 3′ on antisense strands; single underlines, CACCC sites on sense or antisense strands; double underline, GGTGG site; and box, (T)ATAA box; downward arrows demarcate the 350 bp fragment used as a probe for the human ALAD clone. The mouse sequence may be accessed through GenBank 191862.

had 70% of nucleotide sequence similarity with the human DNA (Region 2, Fig. 7). A comparison between the nucleotide sequences from three species was used to reveal conserved and thus functionally important regions of ALAD erythroid-specific promoters. Region 2 also contains the longest stretch of identity (23 bp) among the three species. This stretch contains a double CACCC box, resembling a potential erythroid kruppel-like factor binding site (48). Likewise, two of the three potential GATA-1 sites (asterisks in Fig. 7; –60 to –23 in mouse and rat ALAD, Fig. 4) are identical in all three species. The region between these two GATA-1 sites is identical between mouse and rat and shares 79% identity between mouse and human.

In all regions, the human sequence differs most from the consensus. In region 1, only the human sequence contains an *Alu* repetitive element. Region 3 is exon 1B in mice and rats, whereas in humans, exon 1B is found in region 5, downstream from the position in rodents.

GATA-1 expression during CFU-E maturation

In freshly harvested murine CFU-E, GATA-1 mRNA levels are readily detectable (Fig. 8). During CFU-E maturation, GATA-1 mRNA levels increase 3-fold relative to actin mRNA levels.

ALAD exons 1A and 1B mRNA levels during adult erythropoiesis

After 21 h of CFU-E maturation, both 1A and 1B mRNA levels increase relative to actin (Table 1), which disproved our *a priori* hypothesis that only the tissue-specific, exon 1B, would be regulated during erythropoiesis. Likewise, on two other northern blots probed with ALAD cDNA, increases of ALAD mRNA compared with actin were quantitated to be 2.8 and 3.6 (data not shown). When equal masses of RNA were used in RNase protection assays, the greatest amount of exon 1B was at 21 hours (Fig. 6B). The similarity of results using these two different methods makes it unlikely that the changes observed in ALAD mRNA levels are due solely changes in actin levels during CFU-E maturation. Furthermore, a similar increase in both ALAD exons 1A and 1B occurs during DMSO-induction of MEL cells using an S1-nuclease protection assay (data not shown).

DISCUSSION

The endproduct of the heme biosynthetic pathway, heme, must attain stoichiometric levels that match α- and β-globin production during the substantial increase of hemoglobin accumulation during erythropoiesis. In other words, regulation of a pathway must be coordinated with the regulation of single genes. Heme

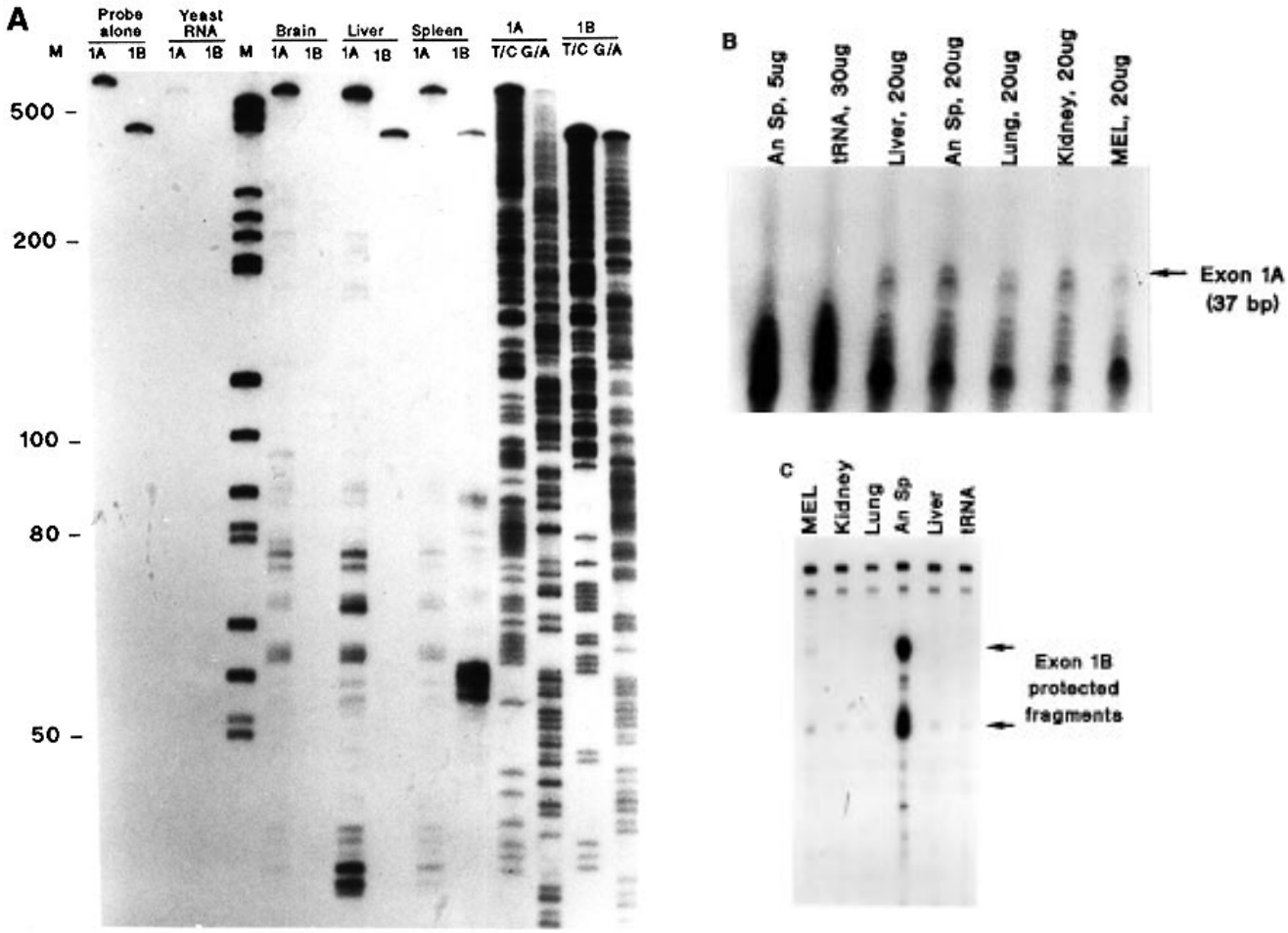

Figure 5. Tissue survey of ALAD exons 1A and 1B using nuclease protection assays. (**A**) S1 nuclease protection assays. Total RNA (50 µg) from brain, liver and spleen were each analyzed with probes specific for either exon 1A or 1B. No bands were visible when yeast RNA was used. Marker fragments (M) were end-labelled pBR322 DNA cleaved with *Hae*III. Maxam and Gilbert (55) sequencing of the probe fragments is shown on the far right. On longer exposure, reannealed probe was visible in all lanes (~500 bp sizes). **(B)** RNase T2 protection assays of exon 1A. Unfractionated RNA from liver, anemic spleen, lung, kidney and uninduced MEL cells were analyzed using the indicated amounts of RNA. **(C)** RNase T2 protection assays of exon 1B using 20 µg total RNA from the tissues indicated. The exon 1B fragments are 112 and 91 bases in length including 9 G residues at the 5′-end which escape nuclease digestion.

synthesis increases ~400-fold during the maturation of CFU-E (S.H. Boyer, unpublished) and, since the pathway consists of eight enzymes (49), the genes encoding the heme synthesis enzymes must be regulated in a precise manner.

Table 1. Quantitation of ALAD alternative first exon utilization during CFU-E maturation.

Time of culture (h)	Exon 1A[a]	Exon 1B[b]
0	1.00	1.00
7	1.28	0.84
14	0.71	0.83
21	3.87	3.04
28	1.77	1.76
35	1.72	1.56

[a]Hybridization ratio of ALAD exon 1A to actin, then normalized to CFU-E at zero time.
[b]Hybridization ratio of ALAD exon 1B to actin, then normalized to CFU-E at zero time.

In the liver, a classical regulatory mechanism is in place: the first enzyme of the heme pathway, ALAS, is rate-limiting and thus, the entire pathway is upregulated when the amount of ALAS increases. On the other hand, the regulation of the pathway in erythroid cells is more complex and it has now been demonstrated that the mRNA level for each of the early heme enzymes increases with different magnitudes during DMSO-induction of MEL cells (50). In cases where the nucleotide sequence of the gene promoter has been reported, there are GATA-1 sites (for ALAS, PBGD, UROD and FC) and NF-E2 sites (for ALAS, PBGD and FC). However, the presence and distance from the transcription start site are very species-specific. For example, in the case of the PBGD, the human erythroid promoter contains an NF-E2 site at –160, but the mouse does not in the same position; the murine PBGD erythroid promoter contains a double CACCC box, whereas the human gene contains only one (51). In this paper, we show that the ALAD gene has a promoter which is utilized only in erythroid cells and that it contains at least two potential GATA-1 binding sites. Furthermore, there is an NF-E2 site located 2.3 kb upstream from the housekeeping promoter, possibly in a distal enhancer. This is reminiscent of β-globin gene

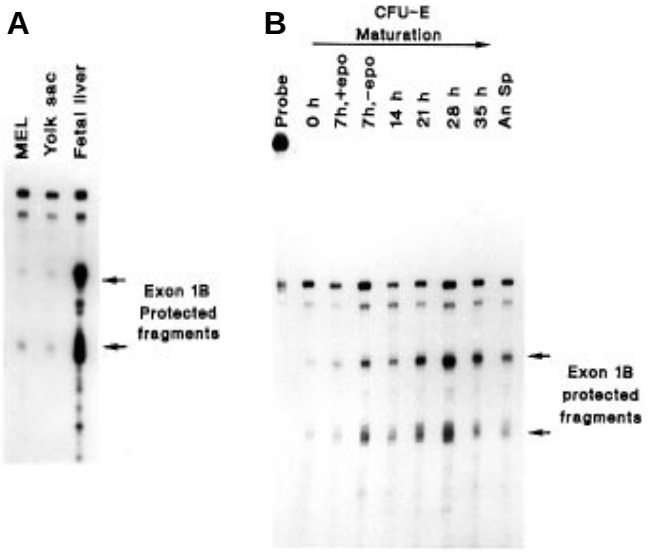

Figure 6. ALAD exon 1B expression during development and differentiation. (**A**) Exon 1B RNase T2 protection assay using 20 µg RNA from uninduced MEL cells (lane 1), 9.5 day murine yolk sacs (lane 2) and 14.5 day fetal livers (lane 3). **(B)** Exon 1B RNase T2 protection assay using 5 µg RNA from CFU-E cultured *in vitro* for the times indicated (lanes 2–8) and compared with 5 µg anemic spleen RNA (An Sp). Except as noted, cultures included 1 U/ml erythropoietin. No significant difference could be measured between cells grown with or without epo. The unhybridized 'probe' (lane 1) shows the total product of the *in vitro* transcription reaction.

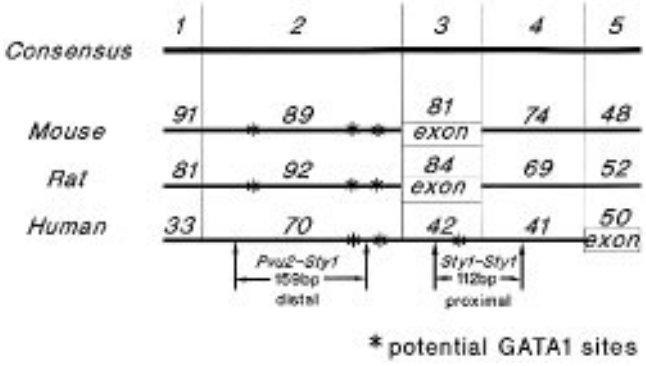

Figure 7. Comparison of mouse, rat and human erythroid-specific regions. The uppermost line represents a 600 bp consensus sequence developed by aligning the best match of human sequence with the 350 bp nucleotide sequence from the mouse, which was used as a probe to isolate the human clone. The position of exon 1B is as described (56). The five regions demarcated are discussed in the text. The numbers above each of the lower three lines are percentages of identical bases with the consensus.

regulation, where binding of NF-E2 to the LCR (along with other factors) opens up the chromatin domain, allowing DNase I hypersensitivity (52) and interaction with transcriptional activators, primarily GATA-1, which bind to the proximal promoter.

Thus, a common theme develops with regards to the type of upstream regulatory elements in the promoters of target genes during erythroid differentiation, but the spatial arrangements of binding sites and therefore, flanking DNA sequences, are distinct, not only for each gene in the pathway, but also in the same gene from different species. Our characterization of the ALAD gene

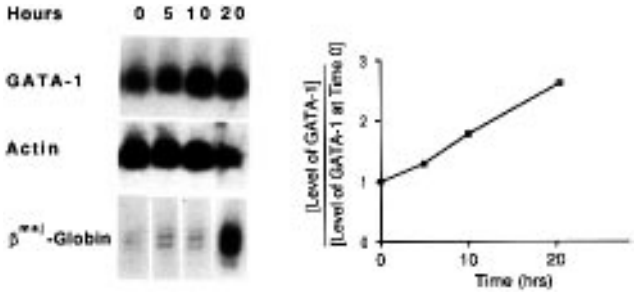

Figure 8. GATA-1 expression during CFU-E maturation. Quantitation of GATA-1 mRNA expression during CFU-E maturation by northern blot analysis. The uppermost panel shows the results of GATA-1 cDNA hybridization to 10 µg total RNA extracted from CFU-E grown for the indicated number of hours. The middle panel shows the results after melting and rehybridizing the filter to a chicken β-actin probe (37). The lowermost panel shows RNase protection by a mouse β^{maj}-globin probe on the same RNA samples as a quality control of the CFU-E preparation. Quantitation of the northern blots is illustrated in the graph.

demonstrates that still another genetic structure has evolved which may be responsive to genetic regulation in an erythroid-specific manner.

While the increase in ALAD mRNA levels during CFU-E maturation measured in this paper was small, it was consistently detected using three different techniques. The relative contribution of exon 1A- and 1B-containing transcripts is difficult to measure precisely, but estimates may be made from both S1-nuclease protection assays and northern blot analyses, in cases where the radiolabeled probes had similar specific activities. In spleen RNA (Fig. 5A), it may be clearly seen that exon 1B transcripts are predominant over exon 1A transcripts and during CFU-E maturation, estimates are that exon 1B transcripts comprise about 2/3 to 3/4 of all ALAD transcripts. This estimation assumes that 1A and 1B probes hybridized with equal efficiency and takes into account the length difference between the two probes.

Measurements of ALAD enzyme levels during CFU-E maturation and calculations based on the known amounts of hemoglobin in BALB/c RBCs, demonstrate that ALAD is not a rate-limiting enzyme of the heme biosynthetic pathway in erythroid cells (S.H. Boyer, unpublished). In fact, ALAD levels are usually 20 times higher than required to participate in the synthesis of the amount of heme eventually realized in the RBC (S. H. Boyer, unpublished). We hypothesized that ALAD might be performing another function besides its role as the second enzyme of the heme biosynthetic pathway and in fact, Guo *et al.* (53) reported that ALAD is the 240 kDa proteasome inhibitor, CF-2. Thus, ALAD acts as an inhibitor of protein degradation via the proteasome. This function may play a critical role in the rapid hemoglobin accumulation observed during erythropoiesis. A third, speculative function of ALAD may be that it is a scavenger of lead, since it has been demonstrated to reversibly bind lead. A better understanding of these multiple roles is needed to completely understand the regulation of the ALAD gene. These multiple ALAD functions may be so critical to erythropoiesis that mutations in the gene are only rarely viable: only four cases of ALAD-deficiencies have been reported in the literature (54).

ACKNOWLEDGEMENTS

We thank Andrea Noyes, Ophelia Rogers and Larry Frelin for their excellent technical assistance, Dave Bodine for his patient

coaching of mouse embryo dissections, Greg Preston for sharing his RNA samples, members of the Kato and Bishop labs for manuscript reviewing and especially, Dr Samuel H. Boyer for his constant encouragement and support. This work was supported by grant DK-38052 from the National Institutes of Health.

REFERENCES

1 Ponka,P. and Schulman,H.M. (1994) *Adv. Exp. Med. Biol.*, **356**, 173–187.
2 Riddle,R.D., Yamamoto,M. and Engel,J.D. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 792–796.
3 Sutherland,G.R., Baker,E., Callen,D.F., Hyland,V.J., May,B.K., Bawden,M.J., Healy,H.M. and Borthwick,I.A. (1988) *Am. J. Hum. Genet.*, **43**, 331–335.
4 Cox,T.C., Bawden,M.J., Abraham,N.G., Bottomley,S.S., May,B.K., Baker,E., Chen,L.Z. and Sutherland,G.R. (1990) *Am. J. Hum. Genet.*, **46**, 107–111.
5 Maguire,D.J., Day,A.R., Borthwick,I.A., Srivastava,G., Wigley,P.L., May,B.K. and Elliott,W.H. (1986) *Nucleic Acids Res.*, **14**, 1379–1391.
6 Cox,T.C., Bawden,M.J., Martin,A. and May,B.K. (1991) *EMBO J.*, **10**, 1891–1902.
7 Schoenhaut,D.S. and Curtis,P.J. (1989) *Nucleic Acids Res.*, **17**, 7013–7028.
8 Fukuda,Y., Fujita, H, Garbaczewski,L. and Sassa,S. (1994) *Blood*, **83**, 1662–1667.
9 Raich,N., Romeo,P.H., Dubart,A., Beaupain,D., Cohen-Solal,M. and Goossens,M. (1986) *Nucleic Acids Res.*, **14**, 5955–5968.
10 Chretien,S., Dubart,A., Beaupain,D., Raich,N., Grandchamp,B., Rosa,J., Goossens,M. and Romeo,P.-H. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6–10.
11 Beaupain,D., Eleouet,J.F. and Romeo,P.-H. (1990) *Nucleic Acids Res.*, **18**, 6509–6515.
12 Mignotte,V., Eleouet,J.F., Raich,N. and Romeo,P.-H. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6548–6552.
13 Romana,M., Dubart,A., Beaupain,D., Chabret,C., Goossens,M. and Romeo,P.-H. (1987) *Nucleic Acids Res.*, **15**, 7343–7356.
14 Kohno,H., Furukawa,T., Yoshinaga,T., Tokunaga,R. and Taketani,S. (1993) *J. Biol. Chem.*, **268**, 21359–21363.
15 Dailey,T.A., Dailey,H.A., Meissner,P. and Prasad,A.R.K. (1995) *Arch. Biochem. Biophys.* **324**, 379–384.
16 Taketani,S., Nakahashi,Y., Osumi,T. and Tokunaga,R. (1990) *J. Biol. Chem.*, **265**, 19377– 19380.
17 Grandchamp,B., Beaumont,C., deVerneuil,H. and Nordmann,Y. (1985) *J. Biol. Chem.*, **260**, 9630–9635.
18 Taketani,S., Inazawa,J., Nakahashi,Y., Abe,T. and Tokunaga,R. (1992) *Eur. J. Biochem.*, **205**, 217–222.
19 Tugores,A., Magness,S.T. and Brenner,D.A. (1994) *J. Biol. Chem.*, **269**, 30789–30797.
20 Ney,P.A., Sorrentino,B.P., McDonagh,K.T. and Nienhuis,A.W. (1990) *Genes Dev.*, **4**, 993–1006.
21 Harrison,P., Rutherford,T., Conkie,D., Affara,N., Sommerville,J., Hissey,P. and Paul,J. (1978) *Cell*, **14**, 61–70.
22 Charnay,P., Treisman,R., Mellon,P., Chao,M., Axel,R. and Maniatis,T. (1984) *Cell*, **38**, 251–263.
23 Boyer,S.H., Landschulz,K.T. and Bishop,T.R. (1987) In Stamatoyannopoulos,G. and Nienhuis,A.W. (eds), *Development Control of Globin Gene Expression.* Alan R. Liss, Inc., New York, pp. 43–54.
24 Landschulz,K.T., Noyes,A.N., Rogers,O. and Boyer,S.H. (1989) *Blood*, **73**, 1476–1486.
25 Nijhof,W. and Wierenga,P.K. (1983) *J. Cell. Biol.*, **96**, 386–392.
26 Landschulz,K.T., Boyer,S.H., Noyes,A.N., Rogers,O.C. and Frelin,L.P. (1992) *Blood*, **79**, 2749–2758.
27 Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) In *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
28 Bishop,T.R., Frelin,L.P. and Boyer,S.H. (1986) *Nucleic Acids Res.*, **14**, 10115.
29 Ullrich,A., Shine,J., Chirgwin,J., Pictet,R., Tischer,E., Rutter,W. and Goodman,H. (1977) *Science*, **196**, 1313–1319.
30 Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
31 Berk,A. and Sharp,P. (1977) *Cell*, **12**, 721–732.
32 Weaver,R. and Weissmann,C. (1979) *Nucleic Acids Res.*, **7**, 1175–1193.
33 Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
34 Myers,R.M., Larin,Z. and Maniatis,T. (1985) *Science*, **230**, 1242–1246.
35 Loh,E., Elliot,J., Cwirla,S., Lanier,L. and Davis,M. (1989) *Science*, **243**, 217–220.
36 Bishop,T.R., Hodes,Z.I., Frelin,L.P. and Boyer,S.H. (1989) *Nucleic Acids Res.*, **17**, 1175.
37 Cleveland,D.W., Lopata,M.A., MacDonald,R.J., Cowan,N.J., Rutter,W.J. and Kirschner,M.W. (1980) *Cell*, **20**, 95–105.
38 Sanger,F., Nicklen,S. and Coulson,A. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
39 Bishop,T.R., Cohen,P.J., Boyer,S.H., Noyes,A.N. and Frelin,L.P. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5568–5572.
40 Fujita,H., Bishop,T.R. and Ishida,N. (1994) In Fujita,H. (ed.), *Regulation of Heme Protein Synthesis.* AlphaMed Press, Dayton, pp. 27–39.
41 Shapiro,M.B. and Senapathy,P. (1987) *Nucleic Acids Res.*, **15**, 7155–7174.
42 Bishop,T.R., Frelin,L.P. and Boyer,S.H. (1990) *Genomics*, **7**, 629–632.
43 Wetmur,J.G., Bishop,D.F., Cantelmo,C. and Desnick,R.J. (1986) *Proc. Natl. Acad. Sci USA*, **83**, 7703–7707.
44 Dynan,W. (1986) *Trends Genet.*, **2**, 196–197.
45 deBoer,E., Antoniou,M., Mignotte,V., Wall,L. and Grosveld,F. (1988) *EMBO J.*, **7**, 4203–4212.
46 Dierks,P., van Ooyen,A., Cochran,M., Dobkin,C., Reiser,J. and Weissmann,C. (1983) *Cell*, **32**, 695–706.
47 Philipsen,S., Talbot,D., Fraser,P. and Grosveld,F. (1990) *EMBO J.*, **9**, 2159–2167.
48 Crossley,M., Tsang,A.P., Bieker,J.J. and Orkin,S.H. (1994) *J. Biol. Chem.*, **269**, 15440– 15444.
49 Kappas,A., Sassa,S., Galbraith,R.A. and Nordman,Y. (1989) In Scriver,C., Beaudet,A.L., Sly,W.S. and Valle,D. (eds), *The Metabolic Basis of Inherited Disease.* McGraw- Hill, New York, pp. 1305–1365.
50 Fujita,H., Yamamoto,M., Yamagami,T., Hayashi,N., Bishop,T.R., DeVerneuil,H., Yoshinaga,T., Shibahara,S., Morimoto,R. and Sassa,S. (1991). *Biochim. Biophys. Acta.*, **1090**, 311–316.
51 Porcher,C., Pitiot,G., Plumb,M., Lowe,S., deVerneuil,H. and Grandchamp,B. (1991) *J. Biol. Chem.*, **266**, 10562–10569.
52 Lowrey,C.H., Bodine,D.M., and Nienhuis,A.W. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1143–1147.
53 Guo,G.G., Gu,M. and Etlinger,J.D. (1994) *J. Biol. Chem.*, **269**, 12399–12402.
54 Ishida,N., Fujita,H., Fukuda,Y., Noguchi,T., Doss,M., Kappas,A. and Sassa,S. (1992) *J. Clin. Invest.* **89**, 1431–1437.
55 Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
56 Kaya,A.H., Plewinska,M., Wong,D.M., Desnick,R.J. and Wetmur,J.G. (1994) *Genomics*, **19**, 242–248.