

Expression of δ -aminolevulinate synthase in avian cells: Separate genes encode erythroid-specific and nonspecific isozymes

(erythroid-specific gene/heme biosynthesis)

ROBERT D. RIDDLE, MASAYUKI YAMAMOTO, AND JAMES DOUGLAS ENGEL*

Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, Evanston, IL 60208

Communicated by David Shemin, October 10, 1988

ABSTRACT A controversy has existed in the literature for the past several years regarding the number of vertebrate genes encoding the mitochondrial protein that initiates the first step in heme biosynthesis, δ -aminolevulinate synthase [ALAS; succinyl-CoA:glycine C-succinyltransferase (decarboxylating), EC 2.3.1.37]. By analysis of chicken ALAS cDNA clones isolated from both liver and erythroid cells, we show that at least two separate genes encode ALAS mRNAs. These experiments show that (i) two different genes encode the ALAS isozymes found in erythroid and in liver tissues, and (ii) while the product of the erythroid gene (ALASE) is expressed exclusively in erythroid cells, the hepatic form of the enzyme is expressed ubiquitously, suggesting that this is the nonspecific form (ALASN) found in all chicken tissues.

δ -Aminolevulinate synthase [ALAS; succinyl-CoA:glycine C-succinyltransferase (decarboxylating), EC 2.3.1.37] is the first enzyme in the heme biosynthetic pathway (1). The mRNA for this mitochondrial enzyme is encoded by one or more nuclear genes, translated into an enzymatically active precursor form, and proteolytically cleaved as it traverses the mitochondrial membrane. The functional form of the enzyme is a homodimer, found in the intracellular compartment of the mitochondrial matrix (2, 3). Although expressed in all tissues, the highest levels of ALAS are found in erythroid and liver cells, where high concentrations of heme are required for hemoglobin or cytochrome P-450 biosynthesis, respectively.

Numerous studies have shown that both the ALAS gene and enzyme are subject to a broad array of regulatory influences in the liver. For example, the enzyme is negatively regulated by hemin (4–11) and can be induced by chemical effectors such as 3,5-dicarbethoxy-1,4-dihydrocollidine (DDC) and allylisopropylacetamide (AIA; refs 12–14). In contrast, ALAS as isolated from reticulocytes appears to be distinct from the hepatic form of the enzyme. The erythroid enzyme is neither induced by the porphyrinogenic effectors of hepatic ALAS nor significantly affected by the addition of hemin (15–18). Furthermore, Bishop *et al.* (19) have partially purified both erythroid and nonerythroid ALAS from guinea pigs and have shown that they differ widely in biochemical properties. Of most direct significance to the studies presented here, Watanabe *et al.* (20) showed immunochemically that ALAS preparations from chicken erythroid and hepatic tissues differ in size, both as precursor and mature proteins. These and other studies (21) have led to the conclusion that the two tissues contain different isoforms of ALAS.

In earlier investigations of the molecular basis of these differences, we made use of an anti-chicken liver ALAS antiserum (which is cross-reactive with the erythroid enzyme; ref. 20) to isolate a partial cDNA clone from a chicken erythroid λ gt11 cDNA library. This recombinant was used in

RNA blotting experiments to show that the ALAS transcript represented by this cDNA clone was present only in erythroid cells (22). The simplest interpretation of those experiments was that erythroid ALAS is encoded by an ALAS gene expressed only in erythroid cells and, consequently, that the erythroid and liver isozymes are transcribed from two different genes. Since that time, other investigators have argued that there is no evidence for an erythroid-specific ALAS gene (23). Using an ALAS cDNA clone isolated from chicken liver (24) to analyze erythroid and hepatic mRNA samples, Elf-erink *et al.* (23) were able to detect only one ALAS mRNA species in either tissue. These data were used to support the conclusions that the form of ALAS as isolated from chicken liver was active in both erythroid and hepatic tissues and that a single gene for ALAS exists in the chicken genome (23).

In this paper, we describe the characterization of cDNAs containing a complete coding sequence for erythroid ALAS. We show that mRNA corresponding to this cDNA is found only in erythroid cells and that this clone is related in sequence to the liver ALAS cDNA (24). Finally, we compare the relative abundance of erythroid and liver ALAS transcripts in both erythroid and hepatic tissue and show that, although hepatic ALAS mRNA is present in erythroid cells, that mRNA is substantially less abundant than the erythroid-specific ALAS mRNA. We conclude, in agreement with our earlier work (22), that at least two different ALAS genes exist in the chicken genome: one that appears to be exclusively expressed in erythroid cells and a second that appears to be expressed in all chicken tissues.

MATERIALS AND METHODS

Erythroid and Liver RNA Isolation. Reticulocyte RNA was isolated from phenylhydrazine-induced anemic chickens as described (25). For hepatic RNA analysis, total cellular RNA was prepared from both induced and normal (untreated) chicken livers. To induce hepatic ALAS transcription, chickens were injected with DDC (100 mg/ml of corn oil per kg of body weight) and then injected 20 hr later with AIA [2% (wt/vol) in saline per kg of body weight]; a second AIA injection was administered 15 hr later. Both induced and untreated animals were sacrificed 5 hr after this final injection. The livers were isolated and perfused with standard saline solution, and RNA was then prepared from the hepatic tissue in the same manner as for erythroid total cellular RNA. RNA samples were poly(A)⁺-selected by using oligo(dT)-cellulose columns as described (26).

Erythroid and Liver cDNA Libraries. The source of erythroid mRNA from which the erythroid cDNA library was prepared was an erythroblast clone transformed by an *erbB*-transducing retrovirus (27). The complexity of the original library was $\approx 5 \times 10^5$ recombinants and was screened

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: ALAS, δ -aminolevulinate synthase; DDC, 3,5-dicarbethoxy-1,4-dihydrocollidine; AIA, allylisopropylacetamide.
*To whom reprint requests should be addressed.

by using modifications of previously described protocols (26). The isolation of the hepatic ALAS cDNA was accomplished by using DDC/AIA-induced liver poly(A)⁺ RNA to prepare double-stranded cDNA that subsequently was used to prepare a phage λ gt11 library as described (28, 29).

DNA Sequence Analysis. cDNA clones were sequenced by conventional dideoxy sequencing protocols (30) for double-stranded plasmids with either Klenow or Sequenase polymerizing enzymes as outlined by the manufacturers (Promega Biotec and United States Biochemical, respectively). All sequencing was performed with priming oligonucleotides complementary to either the phage SP6 or T7 promoters after subcloning of the λ gt11 recombinants into pGEM (Promega Biotec) plasmid vectors. Comparison of chicken erythroid and hepatic ALAS cDNA clones and isozymes was facilitated by use of the IBI/Pustell sequence analysis software (IBI).

RNase Protection. RNase protections were after modifications of Zinn *et al.* (31). pGEM plasmid subclones containing either erythroid or liver ALAS cDNAs (see Fig. 1) were linearized with either *Mst* II or *Mbo* II, respectively, and anti-mRNA sense transcripts were synthesized *in vitro* incorporating [α -³²P]UTP. Approximately 500 pg of radiolabeled transcript was hybridized to each mRNA sample in a final volume of 10 μ l of hybridization solution (80% formamide/40 mM Pipes, pH 6.4/400 mM NaCl/1 mM EDTA) at 43°C overnight; 370 μ l of 0°C RNase solution (10 mM Tris, pH 7.5/5 mM EDTA/300 mM NaCl/100 μ g of RNase A per ml/4 μ g of RNase T1 per ml) was added to the hybridization reactions. The reaction mixtures were incubated at 30°C for 30 min, 10 μ l each of proteinase K at 5 mg/ml and 20% sodium dodecyl sulfate were added, and the samples were incubated for an additional 15 min at 37°C. The solutions were extracted with phenol, precipitated with ethanol, and resuspended in sequencing dye (25). After denaturation at 75°C, the protected RNA fragments were electrophoresed for 1 hr at 20 W on 1-mm vertical 5% polyacrylamide/50% urea gels. The gels were fixed (5% methanol/5% acetic acid), dried, and exposed for autoradiography.

RESULTS

Isolation of Liver and Erythroid ALAS cDNA Recombinants. The previously reported chicken erythroid ALAS cDNA clone (pA4; ref. 22) was used to isolate larger cDNA clones from a chicken erythroid λ gt11 expression library (27) by hybridization screening (26). Recombinant segments of DNA from 20 new cDNA clones were isolated after digestion with *Eco*RI: these corresponded to apparent molecular sizes of 1.2–1.7 kilobase pairs (kbp). These were individually ligated at the *Eco*RI site of pGEM4 to yield the plasmid subclones pAE1–20. Subsequently, 200- to 600-bp fragments of pAE18 (Fig. 1) were further subcloned to facilitate se-

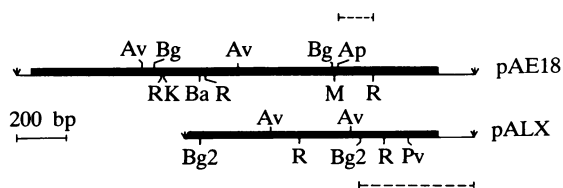


FIG. 1. Structure of the erythroid and liver ALAS cDNAs. (Upper) Restriction map of pAE18, a cDNA clone isolated from the chicken erythroblast cDNA library (27). (Lower) Restriction map of pALX, a cDNA clone isolated from a DDC/AIA-induced chicken liver cDNA library. In both cDNA clones, the 5' end is to the left. The filled box on the line depicts the extent of the coding sequence in either recombinant. Dashed lines adjacent to either sequence depict the specific subclones used in RNase protection experiments (see Fig. 3B). Ap, *Apa* I; Av, *Ava* I; Ba, *Bam*HI; Bg, *Bgl* I; Bg2, *Bgl* II; K, *Kpn* I; M, *Mst* II; Pv, *Pvu* II; R, *Rsa* I; ψ , *Eco*RI linkers.

quence and RNase protection analysis (see below).

To analyze the expression of ALAS mRNAs synthesized in liver cells, we isolated two ALAS cDNA clones from a chicken liver λ gt11 cDNA library, using the previously described anti-chicken liver ALAS antiserum (20). The sequence of the 1.1-kbp cDNA clone pALX (Fig. 1) was identical to the published sequence of chicken liver ALAS cDNA in the region of overlap (ref. 24; data not shown). Fig. 1 shows the restriction maps of both the liver and erythroid cDNA clones used in these analyses.

Sequence of ALAS cDNA Clones Isolated from Erythroid Cell mRNA. Watanabe *et al.* (21) reported the experimentally determined molecular mass of the ALAS precursor protein in chicken erythroid cells to be \approx 55 kDa. The results from complete sequence analysis of the erythroid cDNA clones pAE5 and pAE18 show that a single open reading frame is found that, when translated, is predicted to encode a protein of molecular mass 54,804 Da (Fig. 2). The first methionine codon encountered in this sequence appears to be the appropriate site of translation initiation, since it would initiate a protein of the correct size and since the context of the codon is well matched within the consensus of ideal methionine initiator codons reported by Kozak (32). Watanabe *et al.* (20) also reported that the mature form of chicken erythroid ALAS is \approx 53 kDa; therefore, the 2-kDa mitochondrial signal peptide would correspond to roughly the first 18 amino acid residues of the erythroid ALAS preprotein. The amino acid composition of these first 18 residues (see Fig. 5 below) corresponds well to other reported mitochondrial protein signal sequences in that there is a relative overrepresentation of arginine and leucine (five residues) and a paucity of asparagine, glutamine, valine, and isoleucine (no residues; ref. 33).

Fig. 2 also shows a direct comparison to the DNA sequence of a full-length ALAS cDNA recombinant derived from liver mRNA (24). The two DNA sequences have been aligned to highlight the remarkable lack of nucleotide sequence identity, even where the respective proteins are identical for as many as 16 sequential amino acids (see below). No segment longer than 13 consecutive identical nucleotides is found in DNA sequence comparison of the two maximally aligned proteins.

Erythroid ALAS Is Encoded by a Unique Gene. To examine the relative abundance of the mRNAs that gave rise to cDNA clones pAE18 and pALX (Fig. 1), we initiated RNA blotting experiments. Distinction of the two mRNAs should be readily observable because the liver and erythroid ALAS mRNAs are reported to differ in size (22, 24). Fig. 3A shows duplicate RNA blots, containing identical amounts of either erythroid or (AIA- and DDC-induced) liver poly(A)⁺ mRNA samples, hybridized to either erythroid or liver ALAS cDNA clones (pAE18 and pALX, respectively; Fig. 1). The pAE18 ALAS probe hybridized to a 1.8-kb transcript only in the erythroid mRNA sample, while the pALX ALAS probe hybridized to a 2.2-kb transcript present only in the lane containing liver mRNA. Fig. 3A also shows in lanes 5–8 the same filters after rehybridization to a chicken β -actin cDNA clone (34), which, as anticipated, detected the mature 2-kbp β -actin transcript in both mRNA samples. Clearly, there are comparable amounts of mRNA in both the erythroid and liver samples, and, as reported, the two ALAS mRNAs indeed differ in size by \approx 400 nucleotides (22, 24).

To assess the level of expression of the two forms of chicken ALAS mRNA in liver and erythroid cells more accurately, RNase protection experiments were performed. Radiolabeled antisense transcripts 200 and 1100 nucleotides long, corresponding to either pAE18- or pALX-derived ALAS mRNAs (Fig. 1), respectively, were synthesized *in vitro* and then hybridized to either erythroid or liver poly(A)⁺ RNA. These probes were expected to yield protection fragments of 110 and 420 nucleotides, respectively, after

		10	20	30	40	50	60	70	80	90	100
L 1	CTGTTGCGCT	TCCGCCCGCC	GTGGGGGTGA	CAGCTGCGTG	ACGTCACTTC	CGGTGCGCGG	TAGCTGCGGC	AGGAGGAAGG	ATGAGAGCGG	TGGTGC	CGCGCC
L 101	CTGCCCGTTC	CTGCCCGCGC	TCTCGCAGCC	CTTCCTGCAG	AAGCCCGGGC	CTTCCTTGCT	CTTTTATGCC	CAGCACTGTC	CCAAAATGAT	GGAGCGCGCG	
L 201	CCGCCCGCGC	CCGCCCGAGG	CCTCGCCACA	TCCGCCCGCC	CGGGCAGCA	GGTAGAGGAG	ACCCCTGCGG	CCGACCGCGA	GGCCAAGAAA	GCCAAAGAAC	
E 1											CCCGGCACAG
L 301	TGGCCACGCA	GAACACAGAT	GGGTACACAG	CTCCTGCTGG	CCACCCACCT	GCTGCTGCTG	TCCAGAGCTC	TGCTACAAAA	TGCCCATTC	TGGCAGCTCA	
E 11	GACGATGCGC	GGTATTGTCG	GGTGCCCCCT	CCTGGCCCGA	CACCCCCCCC	TGCGCCCGGC	CTTCGCCACC	GGCGCTGCTG	GCCCTTTTAT	GGGTCTGCGC	
L 401	GATGAACCAC	AAGAGCAGCA	ATGTGTTCTG	CAAAGCCAGC	TTGAACTGTC	AGGAGGATGT	GAAGGAAATG	CAGGTGAGCA	GGAAAGGTAA	AGAATTTGCC	
E 111	CACCCGCGCG	CTCCGGAGCT	GCAGGAAGAC	GTGGAGAGAC	CCCAATCCCC	CGCCGTGGAG	GTTTTGGAGG	AACTGCTGAG	GGACGGCGGG	CGCGCGCTCA	
L 501	AAAATACCAA	CTAATTCCTG	GGTGAGGAAC	ACTGAGGCTG	AGGGAGAAGA	GCAGATGGCC	TTGCTCAAGA	AGTTTAAAGG	TATTATGCTG	AAGCAAAGAC	
E 211	ACAGAACCCT	CGGGACTGTC	ATGGACGAGG	ACGCGTTCCC	CTACGAGGAG	CAGTTCACAG	CGTTCACGT	CGGAGCCCTA	CGCGGACC--	-----	
L 601	CCGAAAGTGT	GTCTCATCTG	CTTCAGGATA	ACTTGCCAAA	ATCTGTATCC	ACCTCCAGT	ATGACCAGTT	CTTTGAGAAA	AAGATAGATG	AAAAGAAGAA	
E 297	----CACACA	TACCGGTAG	TC---ACCGC	TGTGGGGCGG	AGGGCGGAGC	CCCTCCG--	-----TTG	GGCACCCGGC	GAACCGCGCC	CCACACATCC	
L 701	AGATCATACC	TACCGAGTGT	TCAAAACG--	-GTGAACCGA	AAGGCGCAGA	TCTTTCCCAT	CGCAGATGAC	TACTCTGATT	CCCTGATCAC	CAGAAAGAG	
E 381	GTGGAGTTGT	GGTGTCCAG	TGATTACCTC	GGGCTGAGCC	GCCACCCCGC	CGTGTGCGGG	CGCGCCAGGG	CAGCCCTGGA	CGCTCACGGC	CTGGGGCGGG	
L 798	GTGTCTGTGT	GGTGTCCAG	TGATTACCTG	GGCATGAGTC	GTCACCTCCG	TGTGTGCGGA	CGCGTTATGG	ATACACTGAA	ACAACATGTT	GCTGGAGACG	
E 481	GGGTATCCCG	CAATATCGGG	GGGACGTCCC	CCCTGCACGG	CGCCCTGGAG	GGGGCTTGG	CCCTCGTCA	CGGCGCCGCC	TCTTCTGCTC		
L 898	GAGGCACAAG	GAATATCTCA	GGAACAAGCA	AATTTTCATG	CGACTTGGAG	AAAGAACTGG	CTGATCTTCA	TGGAAAAGAT	GCAGCCTTGT	TGTTCTCATC	
E 581	CTGCTTGC	GCCAACGACA	CGCGTTGGA	CACCTTGCC	CGGATCTTAC	CGCGCTGCCA	GGTGTACTCG	GACCGGGGA	ACCAAGCCTC	CATGATTACG	
L 998	TGTCTTGTGA	TCCAATGAT	CCACCCTCTT	CACCTTGTCT	AAAAATGTCG	CAGGTTGTGC	GATCTACTCT	GATCTGTGAA	CATCTGCTC	CATGATCCAG	
E 681	GGCATCCGGC	GCAGGGGGT	CCCCAAATTC	ATCTTCCGTC	ACAACGACCC	CCACCACCTG	GAGCAGCTTT	TGGGGCGCAG	CCCCCCCGG	GTCCCCAAAA	
L1098	GGGATTCGAA	ACAGCAGGGT	GCCAAAACAC	ATCTTCCGCC	ATAACGACGT	CACCATCTT	CGAGAGCTGT	TGAAGAAGTC	TGATCCATCG	ACCCCTAAAA	
E 781	TCGTCCGCTT	CGAATTCGCTG	CACCTCCATGG	ACGGCTCCAT	CGCCCCCTG	CAGCAGCTGT	GTGACGTGGC	TCACGCTTAT	GGGGCGCTGA	CGTTCCCTGGA	
L1198	TTGTTGCGTT	TGAAACTGTG	CACCTCCATGG	ATGGTGTCTG	CTGCCCTCTG	GAAGAGCTGT	GTGATGTGGC	CCACGAGCAC	GGGGCAATCA	CTTTTGTGGA	
E 881	TGAGGTTTAC	GCCCTGGGGC	TCTATGGGGC	CGGGGGCCGA	GGGATCGGGC	AGGGGATGGG	GGTGCAGCAC	AAAGTGGATG	TGTTGTCCCG	CACGCTGGGT	
L1298	TGAAGTGCAT	GCCCTGGGGC	TGTATGGAGC	TGATGGAGC	GGCATAGGGG	ACCCGGATGG	AGTCATGCAC	AAAGTGGATG	TGATCTCTCG	AACGCTCGGG	
E 981	AAAGCATTGG	GGGCCGTGGG	GGGTATCATC	CGGGGGAGCG	AAGCTCTGGT	GGACGCGGTG	CGATCTTTGG	GGCCGGGCTT	CATCTTCAAC	ACGGCTCTGC	
L1398	AAGCCCTTTG	CGTGTGTTGG	AGGATACATC	TCCAGTACAA	GTCGCCCTAT	AGACACTGTC	CGTTCGTATG	CTGCTGGCTT	TATCTTCACA	ACATCCCTGC	
E1081	ACCCACGCC	CCCCATGCTG	GAGAACCTCT	CGCATTAAGT	GTTGGGGAGC	CCCGAGGGC	CGGCCCTAAG	GAGGGCCCA	CAACGGCACG	CCAAAGACTCT	
L1498	CACCCATGCT	CCTGGCTGGT	GCCCTCGAAT	CTGTCCGAAC	TCTGAAAAGT	GCTGAGGGCC	AAGTCTTTAG	GCGCCAGCAC	CAACGCAATG	TGAAGCTCAT	
E1181	GGCGGTCTCA	TTGCGGGATC	GGGGGCTGCC	C-----GCC	CTGCCAGCC	ACATCGTCCC	CGTCAGGTGG	GATGCGGAGG	CC---AACAC	GGCGCTGAGC	
L1598	GAGACAGTGC	CTGATGGATG	CAGGGCTTCC	TGTATGGAGC	TGCGCCAGTC	ACATCATCCC	AATAAGGGTT	GCAGATGCTG	CTAAAATAC	AGAGATCTGT	
E1272	CGCGCGTGC	TGGAGGAGCA	CGGGGTGTAC	GTTACAGCCA	TCAACCACCC	CAGCGTCCCG	CGGGGACAGG	AGCTGTGCTG	GCGCATCGCC	CCCACCCCGC	
L1698	GACAAGTCTG	TGAGCCAAACA	CAGCATCTAT	GTCGCAAGCA	TCAACTACCC	CAGAGTTCTC	CGTGGAGAAG	AGCTGTGCTA	--CGTATTGCT	CCTACACCTC	
E1372	ACCCACGCC	CCCCATGCTG	GAGAACCTCT	CGCATTAAGT	GTTGGGGAGC	CCCGAGGGC	TGGGGCGCAG	CCCGAGGAC	CCCCCGGCTG	CGTCTGCTC	
L1795	ATCACACCCC	TCAAATGATG	AGTTATTTTC	TCGAAAAGCT	GTGGGTCTCA	TGGAAGGATG	TTGGGCTGGA	GCTGAAACCA	CACCTATCAG	CTGAATGCAA	
E1472	ATCGTGTTC	CGCCCTTACC	ACCTCTCCCT	CCTCAGCCCA	CTGAAAAGGG	ATCAGTTTGG	GTTCCGGGGG	GCTCGAGTGT	GG-----	TGACCCCCC	
L1895	CTTCTGAGA	AGACCTTACC	ACTTTGAAGT	GATGAGTAAA	AGGGAAAAGT	CCTACTTCAG	TGGCATGAGC	AAACTATTAT	CTGAGAGTGC	ATGAGAGTAA	
E1563	CACACACCCC	AAAAAAGGAC	CGCCCCCA	AAAAAAGACC	CCCCCAAAA	AAGGGGAC	AACAGCAAAA	AAATATGGGG	GGGGGGAAGA	GTGGAATGCC	
L1995	CAGTGTAAAT	CACTCATATG	CCAATCAGTA	GCATTTTAA	ATTACTTAAT	AAGCATTTTA	ATCATAGTTA	AAGCACTACG	CTCTGAAAAT	AAATTTCTAG	
E1663	AGAAGGAATA	AAATTCGGC									
L2095	AGCCCTGAAA	AAAAAAA									

Fig. 2. DNA sequence comparison of the liver and erythroid cDNAs. Erythroid cell cDNA clones pAE5 and pAE18 were sequenced from double-stranded plasmids as described. pAE5 extended the compiled pAE18 sequence by 45 nucleotides at the 5' end; the two recombinants are identical in every other respect. The sequence of the liver cDNA clone is from ref. 24. Shaded residues represent blocks of identical sequence that are longer than 6 nucleotides. The initiation and termination codons for both proteins are boxed. The left column designation refers to the DNA sequence and nucleotide number for pAE5 and pAE18 ("E") or p105B1 ("L"; ref. 24).

hybridization and RNase treatment of homologous mRNAs. Fig. 3B shows that while the pAE18-derived form of ALAS mRNA was detected exclusively in erythroid cells, a "liver" ALAS transcript was detected in erythroid cell mRNA at greatly reduced abundance ($\approx 1/10$ th) when compared to its relative concentration in normal liver cells. The "liver" transcript (represented by pALX) is far less abundant ($\approx 1/80$ th) than the erythroid transcript in erythroid cells. These data demonstrate both that the pAE18-complementary ALAS mRNA is not expressed in the liver and that pALX-complementary ALAS mRNA is expressed at a low but easily detectable level in erythroid cells. Although these experiments do not qualitatively differ from those reported by Elferink *et al.* (23), we do detect a dramatic quantitative difference in the abundance of pALX-specific mRNA in the two tissues that is not apparent in their experiments.

To verify definitively or to refute the claim that erythroid ALAS is transcribed from a different gene than is liver ALAS mRNA (24), we compared the genetic loci encoding the two mRNAs in genomic DNA blotting experiments, using the cDNA clones pAE18 and pALX. Restriction enzyme-digested chicken chromosomal DNA was hybridized to nick-translated insert segments of either recombinant, and none of the bands detected on the filter that was hybridized to the erythroid cDNA matched the analogous pattern of hybridization to the liver cDNA clone (Fig. 4). These genomic blotting data in conjunction with the RNA analyses (Fig. 3 and ref. 22) argue unequivocally for the existence of a unique erythroid-specific ALAS gene.

Comparison of Erythroid and Hepatic ALAS Amino Acid Sequences. To ascertain the degree of similarity between the two ALAS enzymes, we performed a computer-assisted comparison of the amino acid sequences predicted from the liver and erythroid cDNA sequences (Fig. 5). It is apparent that the amino-terminal 83 amino acids of erythroid ALAS have little, if any, similarity with the corresponding segment of liver ALAS. In addition, the extreme 108 amino-terminal residues of liver ALAS are clearly missing in the erythroid protein. The similarity between the two enzymes begins within the amino-terminal segments of the two proteins, increases, and remains high throughout the remainder of the sequence. In a region spanning the carboxyl-terminal 430 amino acids of the two proteins, erythroid ALAS is 57% identical to liver ALAS; long segments in the internal core of the two enzymes share >75% identity.

DISCUSSION

In this communication we present the sequence of the complete coding region for the erythroid form of the ALAS enzyme. The cDNA clones pAE5 and pAE18 encode an enzyme predicted to be 54,804 Da in molecular mass, a value in excellent agreement with the experimentally determined size of 55 kDa for the precursor erythroid ALAS enzyme as reported by Watanabe *et al.* (20). A comparison of the cDNA clones isolated from erythroid and hepatic cells (24) reveals that the two coding sequences are profoundly divergent at the DNA sequence level (Fig. 2). However, the predicted trans-

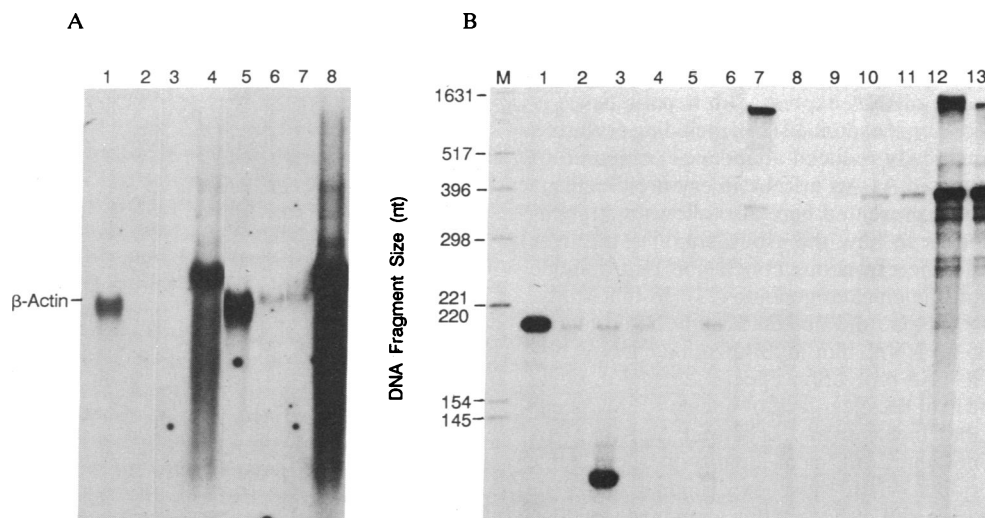


FIG. 3. Expression of ALAS mRNA in chicken erythroid and liver cells. (A) RNA blot hybridization of pAE18 and pALX. Identical samples containing 3 μ g of either anemic hen poly(A)⁺ reticulocyte RNA (odd-numbered lanes) or DDC/AIA-induced liver poly(A)⁺ RNA (even-numbered lanes) were denatured and electrophoresed on a 1.2% agarose/1.1 M formaldehyde gel as described (25) and then blotted to nitrocellulose. The filters were hybridized to nick-translated pAE18 (lanes 1 and 2) or pALX (lanes 3 and 4) and exposed to x-ray film for 3 hr. After the initial exposure, the same filters were rehybridized to a nick-translated β -actin cDNA clone (34), washed, and exposed as before. Lanes 5–8 correspond to lanes 1–4 of the original exposure. (B) RNase protection of liver and erythroid cellular RNAs. Subclones of pAE18 and pALX (graphically depicted by the dashed lines in Fig. 1) were each linearized with restriction enzymes and transcribed with SP6 polymerase to yield initial anti-mRNA sense radiolabeled transcripts (containing both plasmid and recombinant sequence) of \approx 200 and 1100 nucleotides (nt), respectively. Hybridization, RNase treatment, and gel electrophoresis were performed as described. Lanes 1–6 represent an experiment using a pAE18 transcript hybridized to the following poly(A)⁺ RNA samples: none (lane 1), 3.0 μ g of yeast RNA (lane 2), 0.3 μ g of reticulocyte RNA (lane 3), 0.3 μ g of (DDC/AIA-induced) liver RNA (lane 4), 0.3 μ g of normal liver RNA (lane 5), and 3.0 μ g of normal liver RNA (lane 6). Lanes 7–13 represent an experiment using the pALX transcript hybridized to the following poly(A)⁺ RNA samples: none (lane 7), 3.0 μ g of yeast RNA (lane 8), 0.3 μ g of reticulocyte RNA (lane 9), 3.0 μ g of reticulocyte RNA (lane 10), 0.3 μ g of normal liver RNA (lane 11), 3.0 μ g of normal liver RNA (lane 12), and 0.03 μ g of (AIA/DDC-induced) liver RNA (lane 13).

lation products of the two cDNA clones share substantial identity in homology alignment of the two proteins (Fig. 5).

The presumptive tissue specificity of these two mRNAs was addressed by both RNA blotting and RNase protection ex-

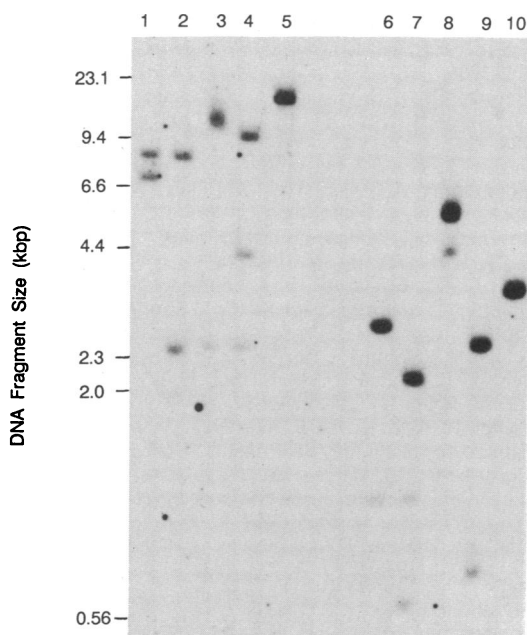


FIG. 4. Genomic DNA blot hybridization of erythroid and liver ALAS cDNAs. Chicken genomic DNA (5 μ g) was digested with restriction enzymes *Pst* I (lanes 1 and 6), *Pst* I/*Bam*HI (lanes 2 and 7), *Bam*HI (lanes 3 and 8), *Bam*HI/*Pvu* II (lanes 4 and 9), or *Pvu* II (lanes 5 and 10); electrophoresed on a 0.8% neutral agarose gel; and blotted to nitrocellulose as described (25, 26). Lanes 1–5 were hybridized to nick-translated pAE18, and lanes 6–10 were hybridized to nick-translated pALX (Fig. 1); the blots were exposed for autoradiography for 60 hr.

L 1	MEAVVRRCPF	LARVSQAFLO	KAGPSLLFYA	QHCPKMEEAA
L 41	PPAAARGLAT	SAARGQQVEE	TPAAQPEAKK	AKEVAQQNTD
E 1				MA AFLRCPLLAR
L 81	GSQPPAGHPP	AAAVQSSATK	CPFLAAQMNH	KSSNVFCKAS
E 13	HPPLARAFAT	GARCPFMGFA	HRAAPELQED	VERPQIPAVE
L121	LELQEDVKEM	QVDRKGFKEFA	KIPTNSVVRN	TEAEGEEQSG
E 53	VLEELLRDGG	AALNRTVRDC	MDEDAFFYEE	QFG-AQLGAL
L161	LLKKFKDIML	KQRPESVSHL	LQDNLPKSVS	TFQYDQFFEK
E 92	RRT-----HT	YRVF-TAVGR	RADAPF---L	GTRGTAPHTS
L201	KIDKKKDHHT	YRVFKT-VNR	KAQIFPMADD	YSDSLITKKE
E123	VELWCSSDYL	GLSRHPAVLR	AARAALDAHG	LGAGGTRNIG
L240	VSVWCSSNDYL	GMSRHRPVCV	AVMDTLKOHG	AGAGGTRNIS
E163	GTSPLHGALE	RALALLHRQP	RAALFSSCPA	ANDTALDTLA
L280	GTSKFHVDLE	KELADLHGKD	AALLFSSCFV	ANDSTLPTLA
E203	RILPGCQVYS	DAGNHASMIQ	GIRRRGVPKF	IFRHNDPHHL
L320	KMLPGCEIYS	DSGNHASMIQ	GIRNSRVPKH	IFRHNVDNHL
E243	EQLLGRSPPG	VPKIVAFESL	HSMDGSIAPL	EELCDVAHAY
L360	RELLKKS DPS	TPKIVAFETV	HSMDGAVCPL	EELCDVAHEH
E283	GALTFVDEVH	AVGLYGARGA	GIAERDGVQH	KVDVVSQTLG
L400	GATTFVDEVH	AVGLYGARGG	GIGDRDGVMH	KMDIISGTLG
E323	KALGAVGGYI	AGSEALVDAV	RSLGPGFIFT	TALPPQRGGG
L440	KAFACVGGYI	SSTSALIDTV	RSYAAGFIFT	TSLPPMLLAG
E363	ALAALQVVG	AEGAALRRAH	QRHAKHLRVL	LRDRGLP--A
L480	ALSVRTLKS	AEGQVLRROH	QRNVKLMRQM	IMDAGLPVVH
E401	LPSHIVEVRW	DAEA-NTRL	RALLEEGLY	VQAINHPTEV
L520	CPSHIIPIRV	ADAANKTEIC	DKLMSQHSIY	VQAINYPTVP
E440	RQOELLRLTA	PTPHHSPPML	ENLADKLESEC	WGAUGLPRED
L560	RGEELL-RIA	PTPHHTPQMM	SYFLEKELAT	WKDVGLLELKP
E480	PPGPS SSSCH	RPLHLSLLSP	LERDQFVGVRG	AAAG
L599	HSSAECNFCR	RPLHFVEMSE	RERSYFSGMS	KLLSVSA

FIG. 5. Comparison of the predicted protein sequences of the erythroid and liver cDNA clones. The predicted sequences of the erythroid and liver-derived cDNA clones (Fig. 2 and ref. 24) are aligned for maximum homology; identical residues are shaded. In the left margin, "E" and "L" refer to the translated sequence of pAE and p105B1 (24) recombinants, respectively.

periments. On the basis of these data, (Fig. 3) and those presented in a previous report (22), we conclude that the erythroid ALAS transcript is expressed exclusively in erythroid cells, whereas the mRNA expressed in hepatic tissue is expressed in every cell type examined (23), including erythroid cells (albeit at dramatically reduced abundance compared to the level of the erythroid ALAS mRNA in erythroid cells).

Analysis of the data presented here also allows a straightforward explanation as to how the group studying the liver ALAS gene came to the erroneous conclusion that a single ALAS gene exists in the chicken genome (23). In that study, a liver ALAS cDNA clone did not detect any heterologous ALAS transcripts in RNA blot hybridization, DNA blot hybridization, or RNase protection experiments. We would propose, on the basis of the clear lack of nucleotide sequence identity (Fig. 2), that the liver and erythroid ALAS cDNA clones simply do not share sufficient identity to allow stable heterologous hybridization in those experiments (23). The enormous quantitative difference in the abundance of liver ALAS mRNA detected in erythroid cells in the experiments reported here and those reported in that same study could be explained if those hybridization experiments were performed in vast cellular RNA excess, instead of using conditions where excess probe was present in the hybridization reaction.

DNA blot analysis demonstrates that the erythroid and liver ALAS cDNA probes detect unique and nonoverlapping chromosomal banding patterns (Fig. 4). These data prove that the two ALAS isozymes are transcribed from separate genomic loci.

We present data here that show that two chicken ALAS proteins are very similar in primary sequence (Fig. 5). This similarity extends relatively evenly throughout the carboxyl-terminal segments of the two proteins but ends abruptly within the amino termini. Of potentially significant interest, the position at which the similarity between the two proteins begins (liver amino acid residue 192; Fig. 5) is very close to the splice acceptor site of the fourth exon of the liver ALAS gene (35). Previous studies of two bacterial ALAS genes have shown that these predicted proteins lack the amino acids corresponding to the first three exons of the chicken liver ALAS gene (36, 37). Furthermore, it has been shown (21, 38) that a degraded form of liver ALAS (corresponding to a molecular mass of ≈ 50 kDa) is fully enzymatically active. These observations, taken together, suggest that the amino termini of the two chicken ALAS enzymes are not necessary for enzymatic activity and that the amino termini of the two isozymes may have evolved independently, perhaps acquiring additional regulatory function(s).

The genomic DNA blot analysis presented here demonstrates that at least two ALAS mRNAs are transcribed from separate genes. We suspect that other mRNA species detected in other chicken tissues (22) might represent new and uncharacterized ALAS (or ALAS-related) transcripts in this gene family. For that reason, we propose the use of a uniform nomenclature for this gene family based on the tissue distribution of those gene products. Since the mRNA from which pAE18 was isolated is expressed exclusively in erythroid cells (Fig. 3), the gene, mRNA, and protein could all be referred to by the designation E (e.g., ALASE for erythroid-specific), whereas since the products of pALX appear to be very widely distributed (23), the gene, mRNA, and protein corresponding to this clone could be designated ALASN (for nonspecific).

We gratefully acknowledge help with and discussion of this work with B. Schafer, D. Endean, and N. Yew. We especially thank Björn Vennström (European Molecular Biology Laboratory) for the gift of the erythroblast cDNA library from which the large ALASE cDNA clones were isolated. This work was supported in part by a post-

doctoral fellowship from the Arthritis Foundation (to M.Y.) and by grants from the National Institutes of Health (to J.D.E.).

- Kikuchi, G., Kumar, A., Talmage, D. & Shemin, D. (1958) *J. Biol. Chem.* **223**, 1214–1219.
- Kikuchi, G. & Hayashi, N. (1981) *Mol. Cell. Biochem.* **37**, 27–41.
- Kikuchi, G. & Hayashi, N. (1982) in *Structure, Dynamics and Biogenesis of Biomembranes*, eds. Sato, R. & Onishi, S. (Japan Sci. Soc. Press, Tokyo), pp. 131–156.
- Sassa, S. & Granick, S. (1970) *Proc. Natl. Acad. Sci. USA* **67**, 517–522.
- Tyrrell, D. L. J. & Marks, G. S. (1972) *Biochem. Pharmacol.* **21**, 2077–2093.
- Yamamoto, M., Hayashi, N. & Kikuchi, G. (1983) *Biochem. Biophys. Res. Commun.* **115**, 225–231.
- Whiting, M. J. (1976) *Biochem. J.* **158**, 391–400.
- Yamamoto, M., Hayashi, N. & Kikuchi, G. (1982) *Biochem. Biophys. Res. Commun.* **105**, 895–900.
- Hayashi, N., Kurashima, Y. & Kikuchi, G. (1972) *Arch. Biochem. Biophys.* **148**, 10–21.
- Yamauchi, K., Hayashi, N. & Kikuchi, G. (1980) *J. Biol. Chem.* **255**, 1746–1751.
- Yamamoto, M., Hayashi, N. & Kikuchi, G. (1981) *Arch. Biochem. Biophys.* **209**, 451–459.
- Granick, S. & Urata, G. (1983) *J. Biol. Chem.* **258**, 821–827.
- Miyakoshi, T. & Kikuchi, G. (1963) *Tohoku J. Exp. Med.* **79**, 199–208.
- Granick, S. (1966) *J. Biol. Chem.* **241**, 1359–1375.
- Wada, O., Sassa, S., Takaku, F., Yano, Y., Urata, G. & Nakao, K. (1967) *Biochim. Biophys. Acta* **148**, 585–587.
- Bottomly, S. & Smithee, G. A. (1968) *Biochim. Biophys. Acta* **159**, 27–37.
- Wood, J. S. & Dixon, R. L. (1972) *Biochem. Pharmacol.* **21**, 1735–1744.
- Wood, J. S. (1974) *Mol. Pharmacol.* **10**, 389–397.
- Bishop, F. D., Kitchen, H. & Wood, W. A. (1981) *Arch. Biochem. Biophys.* **206**, 380–391.
- Watanabe, N., Hayashi, N. & Kikuchi, G. (1983) *Biochem. Biophys. Res. Commun.* **113**, 377–383.
- Yamamoto, M., Fujita, H., Watanabe, N., Hayashi, N. & Kikuchi, G. (1986) *Arch. Biochem. Biophys.* **245**, 76–86.
- Yamamoto, M., Yew, N., Federspiel, M., Dodgson, J., Hayashi, N. & Engel, J. D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3702–3706.
- Elferink, C. J., Srivastava, G., Maguire, D. J., Borthwick, I. A., May, B. K. & Elliott, W. H. (1987) *J. Biol. Chem.* **262**, 3988–3992.
- Borthwick, I. A., Srivastava, G., Day, A. R., Pirola, B. A., Snoswell, M. A., May, B. K. & Elliott, W. H. (1985) *Eur. J. Biochem.* **150**, 481–484.
- Stevens, P. W., Dodgson, J. B. & Engel, J. D. (1987) *Mol. Cell. Biol.* **7**, 1751–1758.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- Zenke, M., Kahn, P., Disela, C., Vennström, B., Leutz, A., Keegan, K., Hayman, M. J., Choi, H.-R., Yew, N. S., Engel, J. D. & Beug, H. (1988) *Cell* **52**, 107–119.
- Young, R. A. & Davis, R. W. (1983) *Science* **222**, 778–782.
- Yamamoto, M., Kure, S., Engel, J. D. & Hiraga, K. (1988) *J. Biol. Chem.* **263**, 15973–15979.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- Zinn, K., DiMaio, D. & Maniatis, T. (1983) *Cell* **34**, 865–879.
- Kozak, M. (1986) *Cell* **44**, 283–292.
- Reid, G. A. (1985) *Curr. Top. Memb. Transp.* **24**, 295–336.
- Cleveland, D. W., Lopata, M. A., MacDonald, R. J., Cowan, N. J., Rutter, W. J. & Kirschner, M. W. (1980) *Cell* **20**, 95–105.
- Maguire, D. J., Day, A. R., Borthwick, I. A., Srivastava, G., Wigley, P. L., May, B. K. & Elliott, W. H. (1986) *Nucleic Acids Res.* **14**, 1379–1391.
- McClung, C. R., Somerville, J. E., Guerinot, M. L. & Chelms, B. K. (1987) *Gene* **54**, 133–139.
- Leong, S. A., Williams, P. H. & Ditta, G. S. (1985) *Nucleic Acids Res.* **13**, 5965–5976.
- Borthwick, I. A., Srivastava, G., Brooker, J. D., May, B. K. & Elliott, W. H. (1983) *Eur. J. Biochem.* **129**, 615–620.