Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 10028–10033, September 1996
Applied Biological Sciences

Organization and chromosomal localization of the gene encoding the mouse acid labile subunit of the insulin-like growth factor binding complex

(150-kDa insulin-like growth factor complex/insulin-like growth factor binding protein/growth hormone)

YVES R. BOISCLAIR*†, DONALD SETO‡, SUSAN HSIEH§, KELLEY R. HURST*, AND GUCK T. OOI§

*Department of Animal Science, Cornell University, Ithaca, NY 14853; ‡DuPont–Merck Pharmaceutical Co., Wilmington, DE 19880; and §Growth and Development Section, Molecular and Cellular Endocrinology Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Communicated by Dale E. Bauman, Cornell University, Ithaca, NY, July 5, 1996 (received for review May 9, 1996)

ABSTRACT **After birth, most of insulin-like growth factor I and II (IGFs) circulate as a ternary complex formed by the association of IGF binding protein 3–IGF complexes with a serum protein called acid-labile subunit (ALS). ALS retains the IGF binding protein–3–IGF complexes in the vascular compartment and extends the $t_{1/2}$ of IGFs in the circulation. Synthesis of ALS occurs mainly in liver after birth and is stimulated by growth hormone. To study the basis for this regulation, we cloned and characterized the mouse ALS gene. Comparison of genomic and cDNA sequences indicated that the gene is composed of two exons separated by a 1126-bp intron. Exon 1 encodes the first 5 amino acids of the signal peptide and contributes the first nucleotide of codon 6. Exon 2 contributes the last 2 nt of codon 6 and encodes the remaining 17 amino acids of the signal peptide as well as the 580 amino acids of the mature protein. The polyadenylylation signal, ATTAAA, is located 241 bp from the termination codon. The cDNA and genomic DNA diverge 16 bp downstream from this signal. Transcription initiation was mapped to 11 sites over a 140-bp TATA-less region. The DNA fragment extending from nt −805 to −11 (ATG, +1) directed basal and growth hormone-regulated expression of a luciferase reporter plasmid in the rat liver cell line H4-II-E. Finally, the ALS gene was mapped to mouse chromosome 17 by fluorescence *in situ* hybridization.**

In vivo, insulin-like growth factor I and II (IGFs) always are complexed to one of a family of six IGF binding proteins (IGFBP-1 to 6) (1, 2). Until birth, binary IGFBP–IGF complexes of 50 kDa predominate in serum, with IGFBP-2 being the most frequently occurring IGFBP moiety (1). In juvenile and adult mammals, however, 80–85% of serum IGFs are found in a ternary complex of 150 kDa composed of one molecule each of IGF, of IGFBP-3, and of a protein that is found only in serum, the acid labile subunit (ALS) (1–3). ALS migrates at an apparent molecular mass of 84–86 kDa by SDS/PAGE, which decreases to 63–66 kDa after N-glycanase treatment (4, 5). Isolated ALS cDNAs predict signal peptide and mature protein of 23 and 580 amino acids in the rat, and of 27 and 578 amino acids in human (5–7).

ALS circulates at 2–3 times the molar concentration of IGFBP-3 and IGFs in humans and rats, driving most of them into ternary complexes (3, 5). Association of IGFBP-3–IGF complexes with ALS has major consequences on the physiology of circulating IGFs. IGFs can no longer leave the vascular compartment, and their $t_{1/2}$ values are extended from 30–90 min to 12–14 h (8, 9). These novel properties are important in the development of the endocrine function of IGFs and in preventing their hypoglycemic potential (1, 2, 9).

ALS mRNAs are detected in the parenchymal cells of liver and in the epithelial cells of the kidney proximal tubule by *in situ* hybridization (10). However, when rat tissues are surveyed by Northern analysis, expression of ALS is detected only in postnatal liver, suggesting that liver is the main site of synthesis (11). Hepatic expression increases rapidly during the first few weeks of postnatal life in rats, and growth hormone (GH) is the most potent hormonal inducer of ALS mRNA in rat liver and in primary hepatocytes (11, 12). Changes in the abundance of mRNA in rat liver are correlated closely with changes in serum ALS (11). Current evidence suggest that circulating levels of ALS in human are subjected to a similar developmental and hormonal regulation (3).

To understand the basis for its unique spatial, developmental, and hormonal regulation, we have cloned the mouse ALS gene, mapped its location to chromosome 17, and identified its promoter. Targeted inactivation of the mouse ALS gene will provide a novel *in vivo* model to study the role of ALS and of the ternary complex in the physiology of circulating IGFs.

MATERIALS AND METHODS

Reagents and General Methods. Restriction endonucleases, polymerases, and DNA modifying enzymes were purchased from Life Technologies (Gaithersburg, MD), New England Biolabs, Pharmacia, and Epicentre Technologies (Madison, WI). Total RNA was extracted from livers of adult rats and mice by the acid guanidium thiocyanate-phenol-chloroform method (13). Double-stranded DNA fragments were labeled with [α-^{32}P]dCTP (3000 Ci/mmol; 1 Ci = 37 GBq) by random priming (13). Sequencing of double-stranded DNA was performed by the dye terminator cycle sequencing method (Perkin–Elmer). Oligodeoxynucleotides used are described in Table 1. When needed, they were labeled with [γ-^{32}P]ATP using T4 DNA polynucleotide kinase (13).

Genomic Cloning. Total RNA from adult rat liver was reversed transcribed with random primers and avian myeloblastosis virus reverse transcriptase (13). A 1.4-kb DNA fragment corresponding to nucleotide (nt) +195 to +1555 of the rat cDNA (ATG, +1) (7) was amplified with *Taq* DNA polymerase using primers R-1S and R-2A (Table 1). Nylon filters (Colony/Plaque Screen; NEN), representing 1×10^6 plaques of a mouse genomic library (129SV in the λ phage Fix

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: IGF, insulin-like growth factor; IGFBP, IGF-binding protein; ALS, acid labile subunit; GH, growth hormone; RACE, rapid amplification of cDNA ends; FISH, fluorescence *in situ* hybridization; TAP–RLPCR, tobacco acid pyrophosphatase–reverse ligation PCR; DAPI, 4′,6-diamidino-2-phenylindole; GAS, interferon-γ-activated sequence; SPI-GLE 1, SPI-GAS-like element 1.

Data deposition: The sequence reported in this paper has been deposited in the GenBank data base (accession no. U66900).

†To whom reprint requests should be addressed.

Table 1. Nucleotide sequence of oligodeoxyribonucleotide used

Name	Position	Sequence
Rat ALS		
R-1S	+195/+217	5′-CACACATCTGCCTGATGACATCC-3′
R-2A	+1533/+1555	5′-GCAAGGAGTTATTCCTGAGGCTG-3′
R-3S	−6/+31	5′-GGAACAATGGCCCTGAGGACAGGAGGCCCAGCCCTGG-3′
R-4A	+2003/+2042	5′-CTTAAATTCAGCGAGCACCTCCCTCCTTTGGGCCCTTGGA-3′
Mouse ALS		
M-5A	+817/+846	5′-CAGGAGGCCAGCCACACGGTTGTGTGACAG-3′
M-6S	+686/+708	5′-CTCTCCGCAGCGTCAAAGCTAAT-3′
M-7A	−2/+32	5′-ACCAGGGCTGGGCTTCCTGTCCTCAGAGCCATTG-3′
M-8A	+81/+97	5′-ATGCTCCAGGATCTGTC-3′
M-9A	+32/+51	5′-CCAGAAAGCCAGAAGCACCA-3′
M-10A	−20/+5	5′-GCCATTGTTCCTTGTACAGAGCTGG-3′

Name, position, and sequence of oligodeoxynucleotides specific to the rat and mouse cDNAs are given. The first letter of the name refers to the species (R for rat, M for mouse), and the last letter refers to the directionality (S for sense, A for antisense). Their position is given relative to the ATG of the respective cDNA.

II; Stratagene), were hybridized with the labeled rat DNA fragment in 5× standard saline phosphate/EDTA (SSPE; 1× SSPE = 20 mM Na phosphate, pH 7.4/180 mM NaCl/1 mM EDTA), 10 mM Na phosphate (pH 6.8), 50% formamide, 100 μg/ml sheared salmon sperm DNA, 10% dextran sulfate, and 2% SDS at 42°C for 16 h. Membranes were washed in 2× SSPE/0.2% SDS at room temperature and in 2× SSPE at 40°C. DNA from positive clones was prepared by polyethylene glycol precipitation, followed by proteinase-K digestion of phage coat and phenol-chloroform extraction (13). The DNA was digested with restriction endonucleases and analyzed by Southern blotting (13). Restriction fragments were isolated from agarose gels, subcloned into the plasmid pGEM7Zf(−) (Promega), and sequenced (13).

cDNA Cloning. The murine ALS cDNA was cloned using the rapid amplification of complementary DNA ends (RACE) method (Marathon; Clontech). Briefly, an oligo(dT) primer was annealed to total RNA from adult mouse liver and extended with Moloney murine leukemia virus reverse transcriptase. The double-stranded cDNAs were ligated to an universal double-stranded adaptor with T4 DNA ligase. Murine ALS cDNAs were amplified for 30 cycles with TaKaRa Ex Taq DNA polymerase (Takara Shuzo; Kyoto) using a primer corresponding to the universal adaptor in conjunction with the mouse primers M-5A (5′ RACE) or M-6S (3′ RACE) (Table 1).

Determination of the Transcription Initiation Sites. Tobacco acid pyrophosphatase (TAP)–reverse ligation PCR (TAP–RLPCR) was used to map the transcription initiation sites as described (14), except that primers and PCR conditions were specific to mouse ALS. Briefly, total RNA from adult mouse liver was treated with DNase I and with calf intestinal alkaline phosphatase. The cap of mRNAs was hydrolyzed with TAP, and the exposed 5′ phosphate ends were ligated with T4 RNA ligase to the RNA linker (GGGCAUAGGCUGACCCUCGCUGAAA). Primer M-8A (1 ng) was annealed to 100 ng of ligated RNA (45 min at 51°C) and extended with Superscript II RNase H⁻ reverse transcriptase (1 h at 51°C). A DNA primer corresponding to the RNA linker (DNA Pr-1, 100 ng), primer M-9A (10 ng), and a master PCR mix were added to the reverse transcription reaction. ALS cDNAs were amplified for 30 cycles (30 sec at 94°C, 30 sec at 59°C, and 30 sec at 74°C) with *Taq* DNA polymerase. Five microliters of the primary PCR reaction was mixed with 5 μl of the master PCR mix and 12.5 ng of primer M-7A or M-10A (added in the ratio of 1:3, labeled/unlabeled) (Table 1), and products were labeled by 5 cycles of linear amplification (30 sec at 94°C, 45 sec at 76°C) using *Taq* DNA polymerase. Mapping of start sites by primer extension was performed exactly as described (13) using total RNA from adult mouse liver and labeled primer M-7A. TAP–RLPCR and primer extension products were resolved on 6% polyacrylamide/7 M urea gels alongside dideoxy chain termination sequencing reactions.

Identification of a Functional Promoter. A genomic DNA fragment corresponding to nt −805 to −11 of the mouse ALS gene (ATG, +1) was subcloned in the sense or antisense orientation into the luciferase reporter gene pGL3-basic (Promega). Plasmids were transfected into the rat hepatoma liver cell line H4-II-E using DEAE-dextran (13). Briefly, cells were grown to 70% confluence in 60-mm dishes with DMEM supplemented with 10% fetal calf serum. Monolayers were washed with Tris-buffered saline and incubated for 15 min with 200 μl of a DNA solution (0.5 mg/ml DEAE-dextran/2 μg luciferase construct/0.05 μg of plasmid pCMV-SEAP in Tris-buffered saline). pCMV-SEAP (Tropix, Bedford, MA) encodes secreted alkaline phosphatase and was used to correct for variation in transfection efficiency. DMEM supplemented with 10% fetal calf serum was added, and the cells were

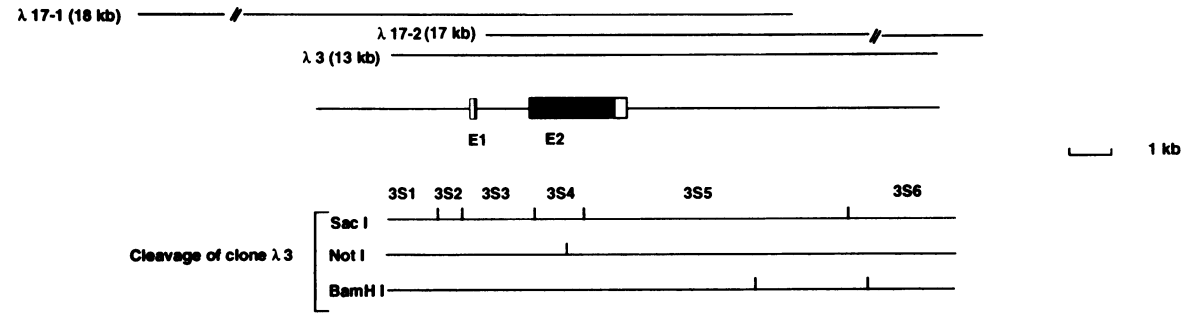

FIG. 1. Organization of the mouse ALS gene. Phage clones containing ALS genomic DNA are depicted at the top, with their length given in parentheses. A schematic map of the mouse ALS gene is shown below the clones. Exons are represented by boxes (coding regions are closed, noncoding regions are open), and intron and flanking regions are represented by horizontal lines. Sites at which the restriction endonucleases *Sac*I, *Not*I, and *Bam*HI cleave the λ3 genomic clone are delineated by vertical bars.

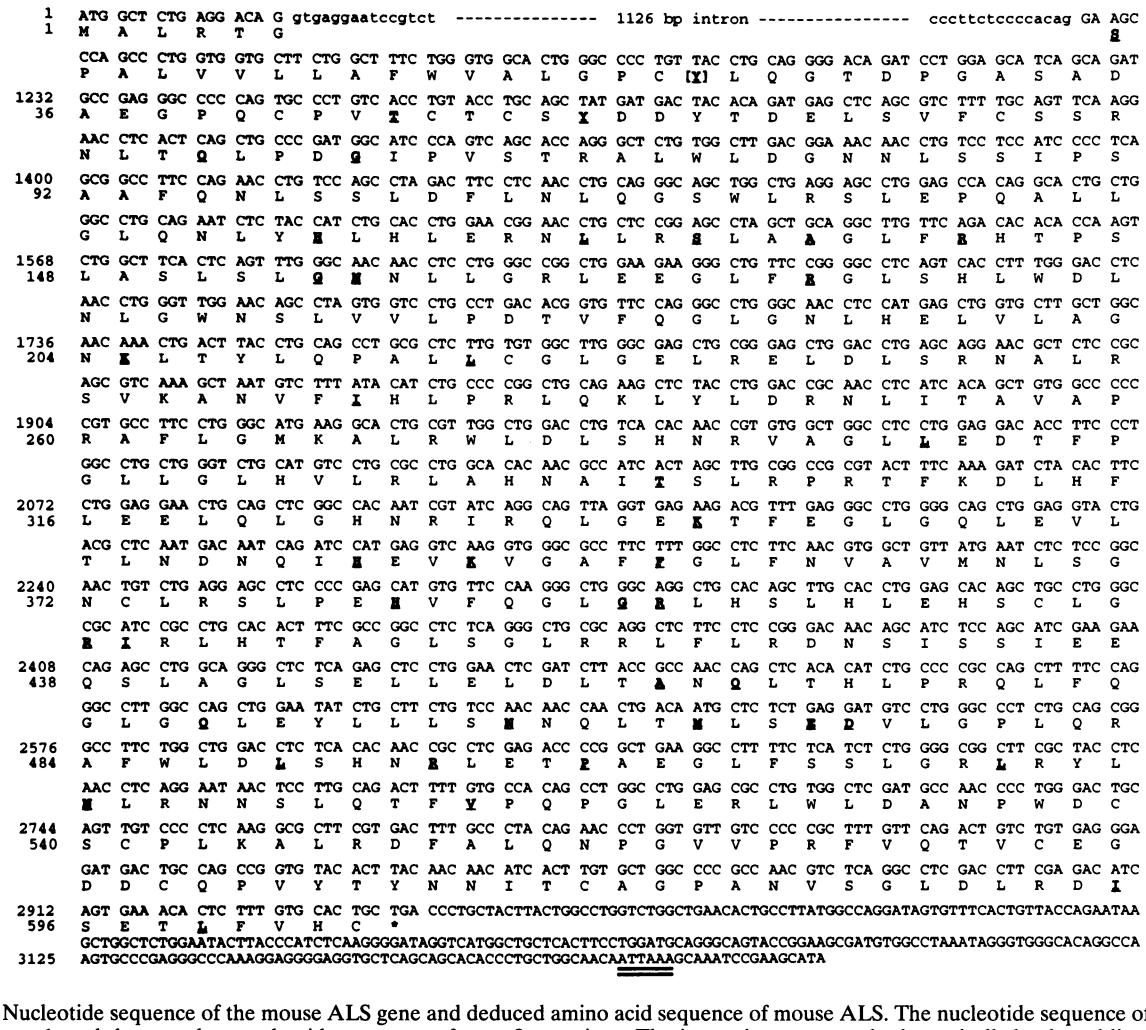

FIG. 2. Nucleotide sequence of the mouse ALS gene and deduced amino acid sequence of mouse ALS. The nucleotide sequence of the coding region of exon 1, and the complete nucleotide sequence of exon 2, are given. The intron is represented schematically by dotted lines except for the 15 nt bordering the exons (written in lowercase letters). Nucleotides are numbered on the left, relative to the translation initiation codon (ATG, +1). The intron/exon boundaries and the 3′ end of exon 2 were determined by comparing genomic and cDNA sequences. The polyadenylylation signal is indicated by a double underline. The deduced amino acids sequence is shown below the nucleotide sequence in standard one-letter abbreviation. Amino acids are numbered on the left relative to the methionine encoded by the initiating ATG. Tyrosine at position 24 is presumed to be the first residue of the mature protein (5, 7) and is bracketed. The stop codon TGA is shown by a star. Amino acids in mouse ALS that differ from rat ALS are indicated by boldface, underlined letters (43 out of 603 amino acids).

returned to the cell culture incubator. After 16 h, medium was changed to serum-free DMEM containing 0.1% bovine serum albumin, and 0 or 100 ng/ml of recombinant human GH (a gift of Genentech). One day later, media were assayed for secreted alkaline phosphatase by chemiluminescence (Phospha-Light; Tropix), and cells lysates were assayed for luciferase activity (15).

Chromosomal Mapping. Fluorescence *in situ* hybridization (FISH) was performed on chromosomes prepared from mouse spleen lymphocytes (16, 17). The genomic insert contained in the phage λ3 (see Fig. 1), biotinylated by nick translation (BioNick Labeling System; Life Technologies), was used as the hybridization probe. The hybridization signal was detected with fluorescein isothiocyanate conjugated to avidin and amplified with biotinylated goat anti-avidin and a second layer of fluorescein isothiocyanate. Chromosomes were identified by staining with 4′,6-diamidino-2-phenylindole (DAPI) (18). Chromosomal assignment of the hybridization signals were obtained by superimposition of FISH signals and DAPI-banded chromosomes.

RESULTS

Cloning of the Mouse ALS Gene. Hybridization of the 1.4-kb rat ALS cDNA probe with the mouse genomic library identified 14 positive clones. To identify clones that encompass the entire ALS gene, DNA from the positive clones was digested with the restriction endonuclease, *Not*I, which released the genomic DNA from the phage and cut the ALS gene once (Fig. 1). Southern blots of the DNA digests were probed with oligonucleotides corresponding to either the 5′ end (R-3S) or the 3′ end (R-4A) of the rat cDNA (Table 1). Three overlapping clones covering ≈28 kb of genomic DNA were identified (Fig. 1).

Phage λ3 was analyzed further as it contains the gene as well as extensive 5′ and 3′ flanking regions. DNA fragments produced by digestion with the restriction endonuclease *Sac*I were probed with the 1.4-kb rat DNA fragment or with the oligonucleotides R-3S or R-4A. These Southern analyses indicated that the entire coding region of the gene is contained within three contiguous fragments 3S3, 3S4, and 3S5 (Fig. 1). These fragments were sequenced and compared with the rat ALS cDNA (7). A region of the 3S3 fragment aligned with the sequence comprised between nt −100 to nt +16 of the rat cDNA. Then, the sequence of the 3S3 fragment diverged for ≈1.1 kb before it matched nt +17 to nt +163 of the rat cDNA. Finally, the 3S4 fragment and the 5′ end of the 3S5 fragment aligned with the rest of the sequence of the rat cDNA. These results suggested that the murine ALS gene is composed of 2 exons and one intron.

To confirm these findings, RACE products covering the entire mouse cDNA were synthesized using total liver RNA.

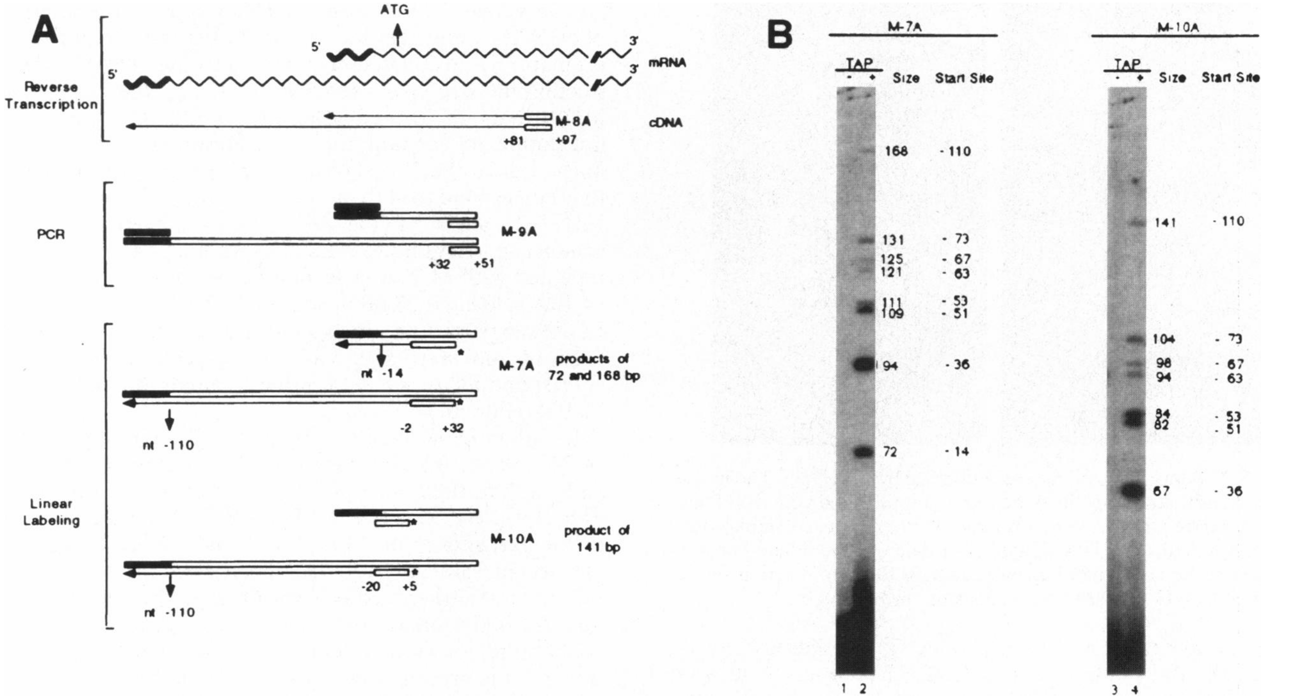

FIG. 3. Identification of the transcription start sites for the mouse ALS gene. (*A*) The TAP–RLPCR assay is illustrated for the shortest and longest mouse ALS mRNA detected in this experiment. ALS primers are represented as open boxes and their positions are given relative to the ATG of the mouse gene. Total RNA from mouse adult liver was treated with DNase I, calf intestinal alkaline phosphatase, and TAP and ligated to the RNA linker (shown as a bold wavy line). Ligated RNA was annealed to primer M-8A and reverse transcribed. ALS cDNAs were amplified with DNA Pr-1 (shown as a solid box) and primer M-9A. Amplified cDNAs were detected by linear labeling with primers M-7A or M-10A labeled with [γ-^{32}P]ATP (shown by asterisks). Transcription initiation sites were deduced from the position of the labeled primers and from the size of the products after subtracting the length contributed by DNA-Pr-1 (25 bp) plus an extra nucleotide to account for the addition of a nontemplated nt at the 3′ end of the amplified DNA. The 3′ end of M-10A is upstream of the cap site of the smallest RNA and therefore cannot detect the corresponding cDNA. (*B*) Transcription start sites detected with primers M-7A or M-10A. The reactions were performed in the absence (−) or in the presence (+) of TAP and analyzed on a 6% polyacrylamide-urea gel. The size of products (on the right of each panel) was obtained by comparing to adjacent dideoxy sequencing reaction. The position of the start site corresponding to each product is given relative to the ATG, +1. Primer M-7A detects all of the start sites. Primer M-10A does not detect the start site at nt −14 because its 3′ end anneals at nt −20 (see *A*).

Primers specific to mouse exon 2 were designed to give RACE products with a 161-nt overlap (M-5A and M-6S, Table 1). The 5′ RACE performed with the antisense primer M-5A yielded a ≈900-bp product, the 3′ RACE performed with the sense

```
                                                                                                          PEA3
-800  CTTTGCACAC AAGGTAGTAG GTTGTTAGGC CCTCCATACT GTGGAGGCCA GCAGAAACTG GCTCTGCCCC TGAGGCTGCG GTGCCCCACT GTCTTCCTGC

                                         HNF-5                                   GAS/SPI-GLE 1   AP-2        Sp1
-700  CAACCCTTTT CTAGGAGGAC TGTGCTGGGG AGAGCAAACA CGAGGTAGCT GGTGGCTCTG CCAGGTGTTC CTAGAAGAGG CCCTGGGGGA GGGGCGGTAG

                                 PEA3             GAS
-600  CCAGCCGTCA GCTGACCTAG AGGTCACAGG AAGACTGAGA CTGGGCCTTA GACAAACCCC TGGACTGGGA GGCACAAATG GTGGGCCCAC TTAGGCCTAA

                                                    PEA3                                   AP-1
-500  GAAACCCCAG GTCAGAGAGC ATTCGAGAAG TATAGGCAAG GCCTTAGGAA ACAGGGCAAA CTGTTTAAGC CTGACAGAGT CATCTGCTCC GTATAGAGGT

                                             GAS
-400  GAGTGGGGGG GTTCCCAGGA ACCCCCGAGC TCCCTCACCT TAGCTGAACC TACTGGGAGA CAGGGCCCCC TACACCTACC TTGAGCACCA ATCCCGAAGG
                                                                 Sp1
                                                                 AP-2                       Sp1
-300  CCTACTCCAC CAAGCCTCTA CCTCCAACAC CAGCAGAGGT GGCTTTCAAT TGCTCCCCCG CCCCTCAATC CCCGAGACTT TGGGGCGGAC AAAGTGCCCC

-200  TCCCCTGCAC CACCCGCATC CAGAGGAGCT TCTGTTGAGA CTTGTTGCCC AGCCCTAGAG GGTTAAAGGA GCCAGAGTGT GGGGCAGGGC CAGCTGAGGG

-100  CTGCCAGCTA CAGGCAGTGG GGAAATCCAC TGGGCAGCAG CCTGTGCTGT AGACGGCCCT TGCTCACTGC CTGCCTGCGG CCAGCTCTGT ACAAGGAACA

   1  ATG
```

FIG. 4. Nucleotide sequence of the 5′ flanking region of the murine ALS gene. Nucleotides are numbered on the left relative to ATG, +1. The sites of transcription initiation are indicated by arrows. Potential cis-elements were identified by searching for homology to the consensus sequences of known transcription factors. They are underlined (thin lines when on the sense strand, thick lines when on the antisense strand) with their names given above their location. Sequence that matched perfectly the consensus recognition sites were found for the transcription factors PEA3 [consensus, AGGAA(G/A)], HNF-5 [T(G/A)TTTG(C/T)], AP2 [CCC(C/A)N(G/C)(G/C)(G/C)], and Sp1 (GGGCGG), and for the interferon-γ-activated sequence (GAS) element (TTNCNNNAA) located between nt −633 to −625. Sequences differing by 1 nt from the consensus site were obtained for the GAS elements located between nt −553 and −545 and between nt −361 and −353, for the SPI-GAS-like element 1 (SPI-GLE 1) element [TTC(C/T)(C/G)(A/T)GAA], and for the transcription factor AP1 [TGA(G/C)T(C/A)A].

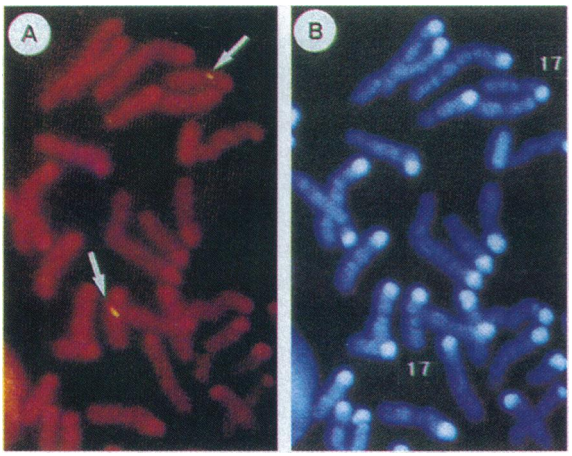

FIG. 5. Mapping of the mouse ALS gene by FISH. (*A*) The mouse genomic insert contained in clone λ3 was biotinylated with dATP and used to probe chromosomes prepared from mouse lymphocytes. Arrows point to the FISH signals on a pair of homologue chromosomes. (*B*) The DAPI banding pattern maps the hybridization signals to region A2–A3 on both homologues of chromosome 17.

primer M-6S yielded a ≈1.4-bp product, giving a mouse cDNA of ≈2.2 kb. Many independent 5′ and 3′ RACE products were sequenced and compared with genomic DNA to determine the structural features of the mouse gene. The ALS gene is comprised of two exons separated by a 1126-bp intron (Figs. 1 and 2). Exon 1 is small (≈150 bp) and encodes the first 5 amino acids of the signal peptide. Splicing occurs after the first nucleotide of codon 6, with the 5′ splice donor and the 3′ splice acceptor conforming to the GT/AG consensus rule (Fig. 2). Exon 2 also encodes the remaining 17 codons of the signal peptide and the 580 amino acids of the mature protein. ATTAAA, the most frequently occurring variant of the canonical polyadenylylation motif AATAAA (19), is present 241 bp downstream of the termination codon. Finally, the cDNA and genomic DNA diverge 16 bp downstream of this motif (Fig. 2). Overall, exon 2 is 2060 bp long and contains 267 bp of non coding sequence at its 3′ end.

The mouse ALS gene encodes a protein that is 92% identical to rat ALS (Fig. 2) and 77% identical to human ALS (6, 7). As with rat and human ALS, mature mouse ALS is organized into 18–20 leucine-rich domains of 24 amino acids (6, 7). Mature mouse ALS contains 13 cysteine residues, 10 of which are clustered outside of the leucine-rich domains, in the COOH and NH_2 terminal domains of the protein. The cysteine residues occur at identical positions in rat ALS (7), and at 12 out of 13 positions in human ALS (6).

Determination of the Transcription Initiation Sites. The 5′ ends of the murine ALS mRNA were mapped first by primer extension. Antisense oligonucleotide M-7A (nt −2 to +32) was annealed to 50 μg of yeast tRNA or 50 μg of total RNA from mouse liver and extended with avian myeloblastosis virus reverse transcriptase. Specific bands corresponding to intitation at nt −125, −110, −92, −73, −63, −53, −51, −36, −29, and −14 were detected (results not shown).

To confirm the authenticity of these sites, we used TAP–RLPCR. Unlike primer extension, TAP–RLPCR does not detect cDNAs that correspond to prematurely terminated extension products (14, 20). Total RNA was prepared from adult mouse liver, treated with TAP, and the RNA linker ligated to the exposed 5′ phosphates. ALS mRNA was reverse transcribed using M-8A (corresponding to nt +81 to +97, Table 1). ALS cDNAs were amplified with DNA Pr-1 and M-9A (nt +32 to +51) for 30 cycles and detected by linear amplification with labeled primer M-7A (nt −2 to +32) (Fig. 3*A*). Products of 168, 131, 125, 121, 111, 109, 94 and 72 bp were detected only when TAP was included in the reaction (Fig. 3*B*, lanes 1 versus 2), indicating that they represent amplification of cDNAs extended to the cap site rather than amplification of prematurely terminated cDNAs or genomic DNA. After subtracting the size of the DNA Pr-1 (25 bp), the nucleotides of M-7A that are located after nt −1 (32 bp), and an extra nucleotide to account for the addition of a nontemplated nucleotide at the 3′ end of the amplified DNA fragment (14), they correspond to start sites at nt −110, −73, −67, −63, −53, −51, −36, and −14 (Fig. 3*B*). To establish further that these products correspond to ALS cDNAs, linear amplification was repeated with M-10A, a second nested primer. The 5′ end of M-10A is located 27 bp upstream of the 5′ end of M-7A (Fig. 3*A*). As expected, products shorter by 27 bp were detected for each of the start site, with the exception of the cDNA corresponding to nt −14, which cannot be detected with M-10A (Fig. 3*B*).

In other experiments, larger faint bands appeared with M-7A and M-10A after long exposure, adding nt −151, −137, and −127 to the family of initiation sites (results not shown). Therefore, the sensitive TAP–RLPCR assay detected additional start sites at nt −151, −137, and −67, and suggest that the primer extension products corresponding to nt −92 and −29 are not authentic start sites. Overall, these assays indicate that transcription of the mouse ALS gene occurs over a ≈140-bp region from at least 11 different sites. None of these start sites is preceded by a TATA element (Fig. 4).

Identification of a Promoter. To test whether the region 5′ to the ATG contained a promoter, a DNA fragment extending from nt −805 to nt −11 was subcloned in the sense or antisense orientation into the promoterless luciferase reporter gene pGL3-basic. These plasmids and pGL3-basic were transfected into the rat H4-II-E cells. In the absence of GH, relative light units obtained were 54,111 ± 4684 for the sense construct, 2702 ± 688 for the antisense construct and 522 ± 14 for pGL3-basic (mean ± SE, $n = 2$). In the presence of 100 ng/ml of GH, luciferase activity of the sense construct increased to 144,000 ± 11,292 relative light unit. Therefore, a GH responsive promoter was present in the nt −805 to −11 DNA fragment. This region contains sites that could be recognized by the transcription factors Sp1, AP1, AP2, HNF-5, and PEA3 (Fig. 4) (21), and sites similar to the GAS element (22) and to the SPI-GLE 1 element (23), which mediate the effects of GH and other cytokines on gene transcription (22, 24).

Chromosomal Location. The chromosomal location of the mouse ALS gene was mapped by FISH using the clone λ3 as the hybridization probe (Fig. 1). The identity of each chromosome was deduced from simultaneous DAPI banding. Of the 100 mitotic figures examined, 90 showed hybridization to chromosome 17 with no hybridization to any other chromosome (Fig. 5). Detailed examination of 10 photographs showed that the hybridization signals mapped within bands A2-A3 of chromosome 17.

DISCUSSION

Using the rat ALS cDNA, we have cloned the mouse ALS gene and completely characterized its structure. The ALS gene spans ≈3.3 kb on mouse chromosome 17 and is composed of two exons interrupted by a 1.126-kb intron. Two features of the gene deserve comments. First, the two exons are very different in terms of coding information. Exon 1 encodes only the first 5 amino acids of the signal peptide. Exon 2 is very long and encodes the entire mature protein of 580 amino acids, 80% of which is made up by 18–20 leucine-rich, repeating domains of 24 amino acids. These repeating units are thought to confer upon ALS its ability to associate with the binary IGFBP-3–IGF complexes (6, 7).

Second, transcription initiation is dispersed over a 140-bp region immediately upstream of the ATG. For the majority of eukaryotic genes, the site of initiation is specified by TATA and/or initiator elements (25). However, none of the transcription initiation sites in the murine ALS gene is preceded by

a TATA element or contained within the loose consensus sequence of an initiator, $PyPyA_{+1}N(T/A)PyPy$ (26). In such genes, other transcription factor binding in proximity to the initiation sites are thought to play a role in positioning the transcription initiation complex. For instance, binding of the transcription factor Sp1 to nearby GC boxes can dictate the initiation site (27, 28). In the mouse ALS promoter, 2 GC boxes that could bind Sp1 are located between nt −243 to −238 and nt −218 to −213, upstream of the region where transcription initiation occurs. It remains to be determined if they are functional and if they convey the information necessary for initiation at only some, or at all of the start sites.

ALS mRNA is expressed at high level only after birth in liver parenchyma and in proximal tubule epithelium of kidney (10, 11). The molecular mechanism by which this pattern of spatial and developmental expression is achieved has not been described, but transcription is the regulated process for many genes displaying tight regulation of expression (29). Computer analysis of the promoter sequence identified a binding site for HNF-5 (hepatic nuclear factor-5), a liver-specific transcription factor (21).

In addition, GH is the most potent hormonal factor stimulating the expression of the ALS gene in postnatal liver (10–12). Most of this stimulation represents an induction of transcription (G.T.O., unpublished data). So far, GH has been shown to regulate the transcription of only a few genes, including c-*fos*, c-*jun*, IGF-I, Spi 2.1, and some of the P450 cytochrome genes (24, 30). Initial events involved in the transmission of the GH signal include dimerization of the receptor by GH, recruitment of the tyrosine kinase, Jak2, to the receptor followed by the phosphorylation of cytoplasmic proteins such as STAT (for signal transducers and activators of transcription), MAP (mitogen-activated protein) kinases, Shc (SH-2 containing protein) and IRS-1 (insulin receptor substrate-1) (24, 30–32). However, current evidence suggest that mediation of the effect of GH at the promoter varies between genes. For the c-*fos* gene, transcriptional activation is mediated by transcription factors bound to the serum response element, perhaps after their phosphorylation by ERK (extracellular regulated kinase) MAP kinases (33). For the Spi 2.1 and the hepatic cytochrome P450 3A10/litocholic acid 6b hydroxylase genes, the action of GH is mediated via STAT5 or STAT5-like proteins binding to sequences resembling the GAS element (23, 34, 35). The proximal promoter region for the ALS gene contains sequences resembling GAS and SPI-GLE 1 elements. Because STAT proteins and MAP kinases are involved in the signaling of many other hormones and growth factors, an unresolved issue remains how GH achieves its specificity (22, 30). Study of the ALS promoter will serve as another paradigm to resolve the mechanisms underlying the stimulation of transcription by GH.

Finally, cloning of the mouse ALS gene is an important step toward our goals of understanding the role of ALS and of the ternary complex in the biology of circulating IGFs. The presently available models, hypophysectomized animal and strains of rodent with defects in the GH axis (9, 11, 36), are limited by their lack of specificity, and by the fact that the other components of the 150-kDa complex are directly (IGF-I) or indirectly (IGFBP-3) regulated by GH (1, 2). Resolution of the mouse gene structure offers the possibility of targeted inactivation of the gene, and the creation of animals in which ALS is absent throughout development (37). Such a model would provide invaluable insights on the role of ALS in the physiology of circulating IGFs.

We thank Dr. M. M. Rechler for his interest and support and for his critical reading of the manuscript. We thank Dr. H. H. Heng for performing the FISH analysis. This work was supported in part by a grant to Y.R.B. from the Cornell Center for Advanced Technology in Biotechnology, which is sponsored by the New York State Science and Technology Foundation and industrial partners.

1. Rechler, M. M. (1993) *Vitam. Horm. (N.Y.)* **47,** 1–114.
2. Jones, J. I. & Clemmons, D. R. (1995) *Endocr. Rev.* **16,** 3–34.
3. Baxter, R. C. (1990) *J. Clin. Endocrinol. Metab.* **70,** 1347–1353.
4. Baxter, R. C., Martin, J. L. & Beniac, V. A. (1989) *J. Biol. Chem.* **264,** 11843–11848.
5. Baxter, R. C. & Dai, J. (1994) *Endocrinology* **134,** 848–852.
6. Leong, S. R., Baxter, R. C., Camerato, T., Dai, J. & Wood, W. I. (1992) *Mol. Endocrinol.* **6,** 870–876.
7. Dai, J. & Baxter, R. C. (1992) *Biochem. Biophys. Res. Commun.* **188,** 304–309.
8. Guler, H.-P., Zapf, J., Schmid, C. & Froesch, E. R. (1989) *Acta Endocrinol. (Copenhagen)* **121,** 753–758.
9. Zapf, J., Hauri, C., Futo, E., Hussain, M., Rutishauser, J., Maack, C. A. & Froesch, E. R. (1995) *J. Clin. Invest.* **95,** 179–186.
10. Chin, E., Zhou, J., Dai, J., Baxter, R. C. & Bondy, C. A. (1994) *Endocrinology* **134,** 2498–2504.
11. Dai, J. & Baxter, R. C. (1994) *Endocrinology* **135,** 2335–2341.
12. Dai, J., Scott, C. D. & Baxter, R. C. (1994) *Endocrinology* **135,** 1066–1072.
13. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. E., Seidman, J. G., Smith, J. A., Struhl, K., Albright, L. M., Coen, D. M., Varki, A. & Janssen, K. (1995) *Current Protocols in Molecular Biology* (Wiley, New York).
14. Boisclair, Y. R. & Brown, A. L. (1995) *DNA Cell Biol.* **14,** 731–739.
15. Boisclair, Y. R., Brown, A. L., Casola, S. & Rechler, M. M. (1993) *J. Biol. Chem.* **268,** 24892–24901.
16. Heng, H. H. Q., Squire, J. & Tsui, L.-C. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9509–9513.
17. Feng, G.-S., Shen, R., Heng, H. H. Q., Tsui, L.-C., Kazlauskas, A. & Pawson, T. (1994) *Oncogene* **9,** 1745–1750.
18. Heng, H. H. Q. & Tsui, L.-C. (1992) *Chromosoma* **102,** 325–332.
19. Wahle, E. & Keller, W. (1992) *Annu. Rev. Biochem.* **61,** 419–40.
20. Fromont-Racine, M., Bertrand, E., Pictet, R. & Grange, T. (1993) *Nucleic Acids Res.* **21,** 1683–1684.
21. Faisst, S. & Meyer, S. (1992) *Nucleic Acids Res.* **20,** 3–26.
22. Darnell, J. E., Jr., Kerr, I. M. & Stark, G. R. (1994) *Science* **264,** 1415–1421.
23. Sliva, D., Wood, T. J. J., Schindler, C., Lobie, P. E. & Norstedt, G. (1994) *J. Biol. Chem.* **269,** 26208–26214.
24. Lobie, P. E., Wood, T. J. J., Sliva, D., Billestrup, N., Waters, M. J., Enberg, B. & Norstedt, G. (1994) *Acta Paediatr. Suppl. (Oslo)* **406,** 39–46.
25. Kollmar, R. & Farnham, P. J. (1993) *Proc. Soc. Exp. Biol. Med.* **203,** 127–139.
26. Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B. & Smale, S. T. (1994) *Mol. Cell. Biol.* **14,** 116–127.
27. Kollmar, R., Sukow, K. A., Sponagle, S. K. & Farnham, P. J. (1994) *J. Biol. Chem.* **269,** 2252–2257.
28. Faber, P. W., van Rooij, H. C. J., Schipper, H. J., Brinkmann, A. O. & Trapman, J. (1993) *J. Biol. Chem.* **268,** 9296–9301.
29. Maniatis, T., Goodbourn, S. & Fischer, J. A. (1987) *Science* **236,** 1237–1245.
30. Rotwein, P., Gronowski, A. M. & Thomas, M. J. (1994) *Horm. Res.* **42,** 170–175.
31. VanderKuur, J., Allevato, G., Billestrup, N., Norstedt, G. & Carter-Su, C. (1995) *J. Biol. Chem.* **270,** 7587–7593.
32. Argetsinger, L. S., Hsu, G. W., Myers, Jr., M. G., Billestrup, N., White, M. F. & Carter-Su, C. (1995) *J. Biol. Chem.* **270,** 14685–14692.
33. Meyer, D. J., Stephenson, E. W., Johnson, L., Cochran, B. & Schwartz, J. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6721–6725.
34. Wood, T. J. J., Sliva, D., Lobie, P. E., Pircher, T. J., Gouilleux, F., Wakao, H., Gustafsson, J. A., Groner, B., Norstedt, G. & Haldosen, L. A. (1995) *J. Biol. Chem.* **16,** 9448–9453.
35. Subramanian, A., Teixeira, J., Wang, J. & Gil, G. (1995) *Mol. Cell. Biol.* **15,** 4672–4682.
36. Gargosky, S. E., Tapanainen, P. & Rosenfeld, R. G. (1994) *Endocrinology* **134,** 2267–2276.
37. Bronson, S. K. & Smithies, O. (1994) *J. Biol. Chem.* **269,** 27155–27158.