Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 5809–5813, August 1988
Biochemistry

Isolation and nucleotide sequence of the cDNA for rat liver serine dehydratase mRNA and structures of the 5′ and 3′ flanking regions of the serine dehydratase gene

(threonine dehydratase/hormonal regulation/consensus sequences)

HIROFUMI OGAWA*†, DUNCAN A. MILLER*, TRACY DUNN*, YEU SU*, JAMES M. BURCHAM‡, CARL PERAINO‡, MOTOJI FUJIOKA†, KAY BABCOCK*, AND HENRY C. PITOT*§

*McArdle Laboratory for Cancer Research, The Medical School, University of Wisconsin, Madison, WI 53706; †Department of Biochemistry, Toyama Medical and Pharmaceutical University, Faculty of Medicine, Sugitani, Toyama 930-01, Japan; and ‡Division of Biological and Medical Research, Argonne National Laboratory, Argonne, IL 60439

Communicated by Van R. Potter, April 15, 1988 (received for review December 29, 1987)

ABSTRACT **Rat serine dehydratase cDNA clones were isolated from a λgt11 cDNA library on the basis of their reactivity with monospecific immunoglobulin to the purified enzyme. Using the cDNA insert from a clone that encoded the serine dehydratase subunit as a probe, additional clones were isolated from the same library by plaque hybridization. Nucleotide sequence analysis of the largest clone obtained showed that it has 1444 base pairs with an open reading frame consisting of 1089 base pairs. The deduced amino acid sequence contained sequences of several portions of the serine dehydratase protein, as determined by Edman degradation. Rat liver serine dehydratase mRNA virtually disappeared from livers of rats fed a protein-free diet for 5 days. Several genomic clones were isolated from two libraries. Determinations of the transcription start site and the structure of the 3′ flanking region of the gene indicated that the coded mRNA is 1504 nucleotides long. The 5′ promoter region contained a variety of sequences similar to several consensus sequences believed to be important for the regulation of specific gene expression.**

L-Serine dehydratase [L-serine hydro-lyase (deaminating), EC 4.2.1.13] catalyzes the α,β-elimination of L-serine to produce pyruvate and ammonia. Studies in several laboratories have shown that the rat liver enzyme also catalyzes the conversion of L-threonine to α-ketobutyrate by a mechanism identical to that converting L-serine to pyruvate (1, 2). The synthesis of the enzyme *in vivo* is enhanced in starvation and diabetes mellitus (3). Serine dehydratase is readily induced by the administration of various hormones that increase gluconeogenesis (4–7), and its synthesis is repressed acutely by the administration of D-glucose (5) and other sugars (8), or chronically by the feeding of diets high in carbohydrates (9). Since our studies were directed to the molecular mechanism of the regulation of the expression of the rat liver serine dehydratase gene, the determination of the structure of the protein, its mRNA, and its gene, especially of regulatory elements of the latter, became essential. Although Noda *et al.* (10) have reported the isolation of a cDNA clone for rat liver serine dehydratase, their clone [1000 base pairs (bp) long] does not represent the entire mRNA, since it was of insufficient size to code for the serine dehydratase subunit that has a molecular weight of 35,000 (10–12).

In this communication, we report the isolation and sequence analysis¶ of a cDNA encoding the entire amino acid sequence of the serine dehydratase subunit. RNA gel blot analysis indicated that the size of serine dehydratase mRNA is close to that of the largest cDNA isolated in this study. The determination of the exact size of DNA complementary to serine dehydratase mRNA was made by S1 nuclease and sequencing of genomic clones of the regions flanking the gene.

MATERIALS AND METHODS

cDNA Cloning. A rat liver cDNA library constructed in λgt11 phage (13) was screened for antibody-reactive plaques as described by Young and Davis (14). Antibodies from an antiserum from a rabbit immunized with the purified serine dehydratase (11) were affinity-purified on a column of serine dehydratase-coupled Sepharose. The nitrocellulose filters (Schleicher & Schuell) blotted with the complex of λgt11-produced chimeric polypeptide and anti-serine dehydratase IgG were incubated with ^{125}I-labeled protein A and they fluorographed at −70°C. The positive clones were plaque-purified, and their cDNA inserts were subcloned into the *Eco*RI site of plasmid pBR322 and used for further characterization. A second screening of the rat cDNA library was performed as above up to the overlaying of the nylon filters (GeneScreen*Plus*, Dupont/NEN). The filters were processed by the method of Benton and Davis (15) with a nick-translated, ^{32}P-labeled SDH1 probe (see text).

Cloning of Genomic DNA. A λ phage Charon 4A library constructed with the *Eco*RI digest of Sprague–Dawley rat liver DNA (16) and a Charon 35 library constructed with the *Sau*3A partial digest of Wistar–Furth rat liver DNA (17) were used in this work. The libraries were kindly supplied by J. Bonner (Phytogen, Pasadena, CA) and C. B. Kasper (University of Wisconsin, Madison, WI), respectively. The two libraries were screened by plaque hybridization with the full-length cDNA as a probe (15).The DNA inserts were subcloned into pBR322 or pGEM-1 (Promega, Madison, WI) for further study.

Other Methods. Hybrid-select translation, RNA gel blotting, and Southern blotting were performed as described (18). Total cellular RNA was isolated by the guanidium thiocyanate/CsCl centrifugation procedure (19) from the livers of rats maintained on either a 0% or 60% (by weight) protein diet for 5 days. Nucleotide sequences were determined by the method of Sanger *et al.* (20). Restriction fragments were ligated to the multiple cloning sites of M13mp18/19. A sequencing kit (Takara Shuzo, Japan) and [α-^{32}P]dCTP (800

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.

§To whom reprint requests should be addressed.

¶The sequences reported in this paper are being deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession nos. J03863 for the cDNA sequence, J03864 for the 5′ flanking sequence, and J03865 for the 3′ flanking sequence).

Ci/mmol; 1 Ci = 37 GBq; Dupont/NEN) were used for sequence reactions, and products were electrophoresed on a continuous 7 M urea/6% polyacrylamide gel (80-cm slab gel). After drying, the gel was subjected to autoradiography with a Fuji RX film for 12–24 hr at room temperature. $NaDodSO_4$/10% polyacrylamide gel electrophoresis was by the method of Laemmli (21) with a calibration kit (product number, MW-SDS-200, Sigma) with *S*-adenosylhomocysteine hydrolase (M_r, 47,300) (22), glycine methyltransferase (M_r, 32,400) (23), and guanidinoacetate methyltransferase (M_r, 26,100) as size references (24).

Amino Acid Sequence Determination. Purified serine dehydratase and the peptides generated by hot $NaDodSO_4$ treatment (25) of the enzyme were analyzed for amino-terminal sequences on an Applied Biosystems model 470A gas-phase sequencer.

Computer Analyses. Protein sequences homologous to serine dehydratase were sought in the National Biomedical Research Foundation data base[‖] with the program of Lipman and Pearson (26).

RESULTS AND DISCUSSION

Isolation of Serine Dehydratase cDNA Clone. In a primary screening of the cDNA library, 5 of 3×10^5 plaques produced relatively strong signals. All of these positive plaques were purified by three successive rounds of screening. The size of the inserts ranged from 0.65 to 1.2 kbp. One of the clones, the 1.2-kbp clone, was designated as SDH1. A hybrid-select translation experiment showed that SDH1 hybridized to the serine dehydratase mRNA (data not shown). Using the SDH1 cDNA as a probe, the same library was rescreened for additional clones. The largest cDNA SDH2 was found to be 1.45 kbp long on agarose gel.

Nucleotide and Predicted Amino Acid Sequences of the cDNA. Digestion with various restriction enzymes showed that SDH1 and SDH2 were identical except that the latter contained an extension of 250 bp in the 5′ end (Fig. 1). The entire nucleotide sequence of SDH2 and the predicted amino acid sequence are presented in Fig. 2. The cDNA consists of 1444 bp with a poly(A) addition signal, AATAAA, at the 3′ end. There are four ATG codons within the first 221 nucleotides, none of which is adequate for initiating an open reading frame capable of encoding a polypeptide having the size of the serine dehydratase subunit reported (10–12). The ATG codon at position 222 initiates an open reading frame of 1089 bp extending to the TGA at position 1311, and this frame encodes a peptide of 363 residues with the calculated molecular weight of 38,392. The sequence around the start site ATG is partly consistent with the consensus sequence, 5′ CCRC-CATGG 3′ (where R is a purine), found by Kozak (27). Since this size was larger than those reported (10–12), we reexamined the subunit molecular weight of the purified serine dehydratase by the method of Laemmli with various globular proteins as size markers. The results clearly indicate that the subunit molecular weight is 38,000 ± 300 (data not shown). When the amino acid compositions of the predicted polypeptide and the purified protein (11) are compared, on the assumption that its molecular weight is 38,000, most of the residues are identical in amount (Table 1).

Since gas-phase Edman degradation failed to release appreciable amounts of phenylthiohydantoin amino acid derivatives, the amino terminus of serine dehydratase is blocked. Hydrolysis of the pure protein with endoproteinase Glu-C (Boehringer Mannheim), which specifically cleaves at the carboxyl site of glutamine residues, was performed, and the largest peptide was isolated by gel electrophoresis and blotted onto glass fiber paper. The partial sequence of this peptide was determined by Edman degradation. As predicted, this sequence begins at the fifth amino acid residue from the first ATG codon found as a plausible start site for translation (Fig. 2).

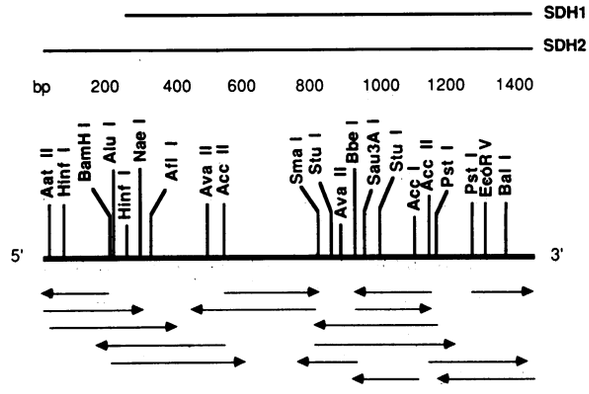

FIG. 1. Partial restriction map of the serine dehydratase cDNA and sequencing strategy. The long lines at the upper portion of the figure show the relative length and overlap of the two characterized cDNA clones, SDH1 and SDH2. The middle portion of the figure shows a partial restriction map of SDH2. The lower portion of the figure shows the direction and the extent to which the sequence of each subclone was determined.

Treatment of serine dehydratase with hot $NaDodSO_4$ generated two peptides with molecular weights of ≈12,000 and ≈25,000. This procedure is known to cleave the peptide bond between aspartic and proline residues (25). Direct application of the peptide mixture to a sequencer revealed the presence of a peptide with the sequence Pro-Leu-Ile-Trp-Glu-Gly-His-Thr-Xaa-Leu-Val-Lys-Glu-Leu-Lys-Glu-. This sequence is identical with the cDNA-derived sequence from residue 175 to residue 190. An aspartic residue is present in the predicted sequence at the position preceding the amino-terminal proline residue. No other Asp-Pro bond occurs in the predicted sequence, indicating that the SDH2 clone is complementary to serine dehydratase mRNA.

Dietary Regulation of Serine Dehydratase mRNA. Total cellular RNAs isolated from the livers of rats maintained on a 0% and 60% protein diet were denatured and fractionated on an agarose gel containing formaldehyde. Hybridization was carried out with a nick-translated ^{32}P-labeled SDH2 probe. As seen in Fig. 3, there is essentially no hybridizable RNA in the livers of rats maintained on a protein-free diet, whereas a substantial amount of serine dehydratase mRNA is present in the livers of animals fed a 60% protein diet. The levels of albumin mRNA as determined by probing with a ^{32}P-labeled albumin cDNA were unaffected by changes in dietary protein. The serine dehydratase mRNA appears to be composed of a single component of ≈1500 nucleotides. This value is in agreement with that reported by Noda *et al.* (10).

Cloning of Serine Dehydratase Genomic DNA. A number of genomic clones for serine dehydratase were isolated from two λgt11 genomic libraries. Of these, one clone in Charon 35 had fragments of 3, 4.2, and 6.6 kbp, and one clone from Charon 4A contained fragments of 3 and 14 kbp. Southern blot analysis indicated that the 3-, 6.6-, and 14-kbp fragments hybridized to SDH2. The 6.6- and 14-kbp fragments overlapped in the 5′ region.

Structure of Regions Flanking the Gene. In Fig. 4 are seen the nucleotide sequences of the 5′ and 3′ flanking regions of the serine dehydratase gene. The transcription start site (cap site) was determined by S1 nuclease mapping (33) (data not

[‖]Protein Identification Resource (1987) Protein Sequence Database (Natl. Biomed. Res. Found., Washington, DC).

```
                  CCCTCTAGATCAGGACGTCGCCGGGGTGGCTGTGACTTGGCCAAGTGCTCGCATGAGTCAAAT     63

GACAAGGAAGAGACTTCTGCCGTGGAACCCATGCCGCACCGGCCACCTTTGCCAAGACCGCCTGTGCCTTTTTCTCTCG    142

CAGGTGCGGCGGGGCATACCTGTGATCCCAGCAATTGGGAGACTGAGACAGGAGGATCCAACCTTCAAAGCTACATGCC    221

ATG GCT GCC CAG GAG TCC CTG CAC GTG AAG ACC CCA CTA CGT GAC ACG ATG GCA TTG TCC    281
Met Ala Ala Gln Glu Ser Leu His Val Lys Thr Pro Leu Arg Asp Ser Met Ala Leu Ser
1                                   10                                      20
AAA GTG GCC GGC ACT AGT GTG TTC CTT AAG ATG GAC AGC TCT CAG CCC TCT GGC TCC TTC    341
Lys Val Ala Gly Thr Ser Val Phe Leu Lys Met Asp Ser Ser Gln Pro Ser Gly Ser Phe
                                    30                                      40
AAG ATC CGA GGC ATT GGG CAT CTC TGC AAG ATG AAG GCA AAA CAA GGC TGT AAA CAT TTC    401
Lys Ile Arg Gly Ile Gly His Leu Cys Lys Met Lys Ala Lys Gln Gly Cys Lys His Phe
                                    50                                      60
GTC TGC TCT TCA GTC GTC CAG ATT TGG GGT TCC AGA ATG AGG GGC AGA AGT CAC TCT GGA    461
Val Cys Ser Ser Val Val Gln Ile Trp Gly Ser Arg Met Arg Gly Arg Ser His Ser Gly
                                    70                                      80
GAT GAG CAG CCC CAC GTG AGG TCC CAG GCC CTC CTT CCT GAT ACA CCC TCT CCA CTG ACA    521
Asp Glu Gln Pro His Val Arg Ser Gln Ala Leu Leu Pro Asp Thr Pro Ser Pro Leu Thr
                                    90                                     100
GCG GGC AAC GCG GGC ATG GCG ACT GCC TAT GCT GCC AGG AGG CTG GGC CTC CCA GCC ACT    581
Ala Gly Asn Ala Gly Met Ala Thr Ala Tyr Ala Ala Arg Arg Leu Gly Leu Pro Ala Thr
                                    110                                    120
ATT GTT GTG CCA AGC ACC ACA CCT GCC CTC ACC ATT GAG CGG CTG AAG AAC GAA GGG GCC    641
Ile Val Val Pro Ser Thr Thr Pro Ala Leu Thr Ile Glu Arg Leu Lys Asn Glu Gly Ala
                                    130                                    140
ACA GTT GAA GTG GTG GGA GAG ATG CTG GAT GAG GCC ATC CAA CTG GCC AAG GCT CTG GAA    701
Thr Val Glu Val Val Gly Glu Met Leu Asp Glu Ala Ile Gln Leu Ala Lys Ala Leu Glu
                                    150                                    160
AAG AAC AAC CAA GGT TGG GTG TAC ATC TCC CCC TTC GAT GAC CCT CTC ATC TGG GAA GGC    761
Lys Asn Asn Pro Gly Trp Val Tyr Ile Ser Pro Phe Asp Asp Pro Leu Ile Trp Glu Gly
                                    170                                    180
CAC ACT TCC CTT GTG AAG GAG CTG AAG GAG ACA CTG AGC GCC AAG CCC GGG GCC ATT GTG    821
His Thr Ser Leu Val Lys Glu Leu Lys Glu Thr Leu Ser Ala Lys Pro Gly Ala Ile Val
                                    190                                    200
CTG TCT GTG GGC GGT GGA GGC CTG CTG TGC GGA GTG GTC CAG GGG CTG CGG GAG GTG GGC    881
Leu Ser Val Gly Gly Gly Gly Leu Leu Cys Gly Val Val Gln Gly Leu Arg Glu Val Gly
                                    210                                    220
TGG GAG GAT GTG CCC ATC ATC GCC ATG GAG ACC TTC GGC GCC CAC AGC TTC CAC GCT GCC    941
Trp Glu Asp Val Pro Ile Ile Ala Met Glu Thr Phe Gly Ala His Ser Phe His Ala Ala
                                    230                                    240
GTC AAG GAA GGA AAG CTG GTC ACC CTG CCC AAG ATC ACC AGT GTT GCC AAG GCC TTG GGT   1001
Val Lys Glu Gly Lys Leu Val Thr Leu Pro Lys Ile Thr Ser Val Ala Lys Ala Leu Gly
                                    250                                    260
GTG AAC ACT GTG GGG GCA CAG ACC CTG AAG CTG TTT TAC GAA CAC CCC ATT TTC TCT GAG   1061
Val Asn Thr Val Gly Ala Gln Thr Leu Lys Leu Phe Tyr Glu His Pro Ile Phe Ser Glu
                                    270                                    280
GTC ATC TCA GAC CAG GAG GCT GTG ACT GCT ATC GAG AAG TTC GTA GAC GAT GAG AAG ATC   1121
Val Ile Ser Asp Gln Glu Ala Val Thr Ala Ile Glu Lys Phe Val Asp Asp Glu Lys Ile
                                    290                                    300
CTG GTG GAG CCC GCG TGT GGC GCT GCC CTG GCT GCA GTG TAC AGC GGT GTG GTG TGC AGG   1181
Leu Val Glu Pro Ala Cys Gly Ala Ala Leu Ala Ala Val Tyr Ser Gly Val Val Cys Arg
                                    310                                    320
CTG CAG GCT GAG GGC CGA CTG CAA ACC CCA CTG GCC TCG CTG GTT GTC ATT GTG TGT GGT   1241
Leu Gln Ala Glu Gly Arg Leu Gln Thr Pro Leu Ala Ser Leu Val Val Ile Val Cys Gly
                                    330                                    340
GGC AGC AAC ATC AGC CTG GCA CAG CTG CAG GCA CTC AAG GCA CAG CTG GGC CTG AAT GAG   1301
Gly Ser Asn Ile Ser Leu Ala Gln Leu Gln Ala Leu Lys Ala Gln Leu Gly Leu Asn Glu
                                    350                                    360
CTA CTC AAG TGA TATCTGCTGCTGCCCTGGCCACCCTGAGGGGTCACCAGCACCCCTGAGTAGGCTGGGTGGGCG   1376
Leu Leu Lys

TCCGCCTGACAGTGGCCCACCCTCCTTTATCCATGTTTATAATATGCACTTTTTCATTGTAAATAAAA                1444
```

FIG. 2. Nucleotide sequence of SDH2 and deduced amino acid sequence. The nucleotide sequence numbers are shown at the right margin, and amino acid sequence numbers are shown below the amino acids. The methionine corresponding to ATG at position 222 is numbered 1. The amino acid sequences determined by Edman degradation are underlined, and the poly(A) addition signal is boxed.

shown) and is numbered +1. The sequence around the cap site is similar to the general observation that pyrimidines are located at positions −5, −1, and +2 to +6 (34). In the 5′ flanking region, there are no typical "TATA" and "CAAT" boxes, which are usually located ≈30 and ≈70 nucleotides, respectively, upstream from the cap site (35). Instead, its variations, "ATA" (36) and "CAT" boxes (23), occur at positions −194 and −222, respectively. Within this region are the sequence CTTTCCAC, which is complementary to the enhancer core of a 72-bp repeat sequence of simian virus 40 (37), and the sequence TCTCACCCCA, which is homologous to the enhancer core of polyoma virus (38). Upstream of the ATA box, there are a number of "GC"-like boxes with homology to those reported to be the binding site of transcription factor Sp1 (34). The sequence ATTTGCAA at position −357 is almost identical to that of lymphoid enhancer motif (39). There are 19 sets of alternating TG sequences near position −1200, which might favor a change of chromatin structure (30). There exist two putative glucocorticoid response elements (32) at positions −289 and −2241. Glucocorticoids regulate serine dehydratase *in vivo* or in primary cultures of hepatocytes (4–7, 11, 12). Jantzen *et al.* (29) have reported that the two glucocorticoid response elements at positions −2472 and −2495 in the rat tyrosine aminotransferase gene function as a transcriptional enhancer. By analogy the far upstream sequence in the serine dehydratase gene may also be important for gene expression.

In the 3′ flanking sequence of the gene, a poly(A) addition signal, AATAAA (28), occurs at position 124. Because both of the cDNA clones (SDH1 and SDH2) examined lack poly(A) tails, we cannot assign the poly(A) addition site in the cDNA sequence. However, we assume the cytidine residue at position 155 to be the site of poly(A) addition, since (*i*) the boundary C′AA sequence [where the prime indicates the poly(A) addition site] is conserved in most eukaryotic genes (40) and (*ii*) the distance between this residue and the

Table 1. Amino acid composition of serine dehydratase

Amino acid	Number of residues per monomer	
	Predicted from cDNA sequence	Amino acid analysis of purified protein
Aspartic acid	11	18
Asparagine	7	
Glutamic acid	23	39
Glutamine	15	
Half cystine	7	7
Serine	28	25
Glycine	34	34
Histidine	9	7
Arginine	12	9
Threonine	19	19
Alanine	39	42
Proline	17	17
Tyrosine	4	4
Valine	35	35
Methionine	8	6
Isoleucine	18	18
Leucine	41	41
Phenylalanine	9	10
Lysine	23	24
Tryptophan	4	4
Total	363	359

Values described in ref. 11 are corrected by multiplying by 1.11, on the basis that the subunit molecular weight of serine dehydratase is 38,000.

AATAAA sequence (26 nucleotides) is in the range of 10–30 nucleotides reported for other mRNA 3′ sequences (28). If the postulated residue is indeed the site of addition, the length of DNA complementary to serine dehydratase mRNA is 1504 bp. This value approximates the size of the mRNA determined by RNA gel blot hybridizations of total cell RNA (Fig. 4) and independently by Noda *et al.* (10). These results suggest that the poly(A) tail of serine dehydratase mRNA may be relatively short. This is of interest since one of the functions of the poly(A) tail on mRNAs is to enhance stability of the mRNA *in vivo* (40–43). We find an A + T-rich sequence in the 3′ end of the gene (84% abundance from positions 98 to 153). Shaw and Kamen (44) reported that a highly conserved 50–60 nucleotide A + T-rich sequence found in the 3′ noncoding region of mRNAs may be involved in mRNA stability. Since the stability of serine dehydratase in RNA appears to vary in normal and neoplastic rat liver (45), further investigations will be focused on elucidating the mechanisms of its regulation in these tissues.

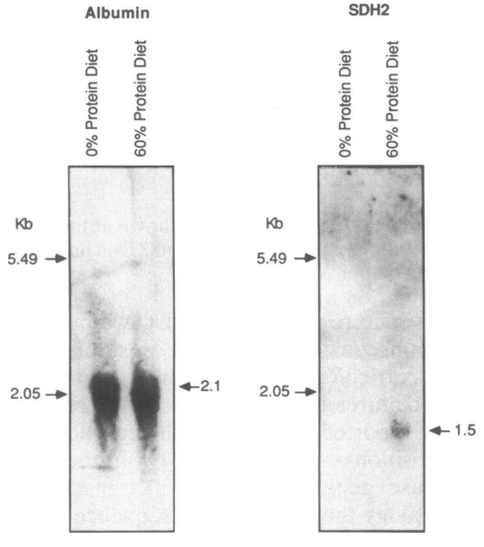

FIG. 3. RNA gel blot analysis of mRNAs for albumin and serine dehydratase under various dietary conditions. Total cellular RNA was prepared. Ten micrograms of various RNAs pretreated with formaldehyde/formamide was electrophoresed in a 1.5% agarose gel, blotted on a nylon filter, and hybridized with the ^{32}P-labeled probe (20). (*Left*) Steady-state level of albumin mRNA from the livers of rats fed 0% protein diet or from the livers of rats fed 60% protein diet. (*Right*) Steady-state level of serine dehydratase mRNA from the livers of rats fed 0% protein diet or from the livers of rats fed 60% protein diet. Ribosomal RNAs were used as size markers.

This work was supported by grants from the National Cancer Institute (CA-07175 and CA-22484) and a grant from the National Science Foundation (DMB8514305) to the University of Wisconsin Biotechnology Center. We express our sincere appreciation to Ms. Susan Moran for expert technical assistance, to Mrs. Kristen Adler and Mrs. Mary Jo Markham for expert technical typing, and to Dr. Ilse Riegel for many editorial comments on the manuscript.

1. Nakagawa, H. & Kimura, H. (1969) *J. Biochem. (Tokyo)* **66,** 669–683.
2. Goldstein, L., Knox, W. E. & Behrman, E. J. (1962) *J. Biol. Chem.* **237,** 2855–2860.
3. Ishikawa, E., Ninagawa, T. & Suda, M. (1965) *J. Biochem.* **57,** 566–513.
4. Nakagawa, H. (1971) *Seikagaku* **43,** 1–22.
5. Peraino, C. & Pitot, H. C. (1964) *J. Biol. Chem.* **239,** 4308–4314.
6. Inoue, H. & Pitot, H. C. (1970) *Adv. Enzyme Regulat.* **8,** 289–296.
7. Morris, J. E. & Peraino, C. (1976) *J. Biol. Chem.* **251,** 2571–2578.
8. Straub, K., Jenniseen, H. & Kroger, H. (1970) *Z. Gesamte Exp. Med.* **152,** 294–305.
9. Ogura, M. (1969) *Agric. Biol. Chem.* **30,** 1054–1057.
10. Noda, C., Tomomura, M., Nakamura, T. & Ichihara, A. (1985) *Biochem. Biophys. Res. Commun.* **132,** 232–239.
11. Inoue, H., Kasper, C. B. & Pitot, H. C. (1971) *J. Biol. Chem.* **246,** 2626–2632.
12. Noda, C., Tomomura, M., Nakamura, T. & Ichihara, A. (1984) *J. Biochem. (Tokyo)* **95,** 37–45.
13. Mueckler, M. M. & Pitot, H. C. (1985) *J. Biol. Chem.* **260,** 12993–12997.
14. Young, R. A. & Davis, R. W. (1983) *Science* **222,** 778–782.
15. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
16. Sargent, T. D., Wu, J.-R., Sala-Trepat, J. M., Wallace, R. B., Reyes, A. A. & Bonner, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3256–3260.
17. Falany, C. N., McQuiddy, P. & Kasper, C. B. (1987) *J. Biol. Chem.* **262,** 5924–5930.
18. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
19. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
20. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
21. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
22. Ogawa, H., Gomi, T., Mueckler, M. M., Fujioka, M., Backlund, P. S., Jr., Aksamit, R. R., Unson, C. G. & Cantoni, G. L. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 719–723.
23. Ogawa, H., Konishi, K., Takata, Y., Nakashima, H. & Fujioka, M. (1987) *Eur. J. Biochem.* **168,** 141–151.
24. Ogawa, H., Date, T., Gomi, T., Konishi, K., Pitot, H. C., Cantoni, G. L. & Fujioka, M. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 694–698.
25. Rittenhouse, J. & Marcus, F. (1984) *Anal. Biochem.* **138,** 442–448.
26. Lipman, D. J. & Pearson, W. R. (1985) *Science* **216,** 1435–1441.
27. Kozak, M. (1986) *Cell* **44,** 283–292.
28. Nevin, J. R. (1983) *Annu. Rev. Biochem.* **52,** 441–466.
29. Jantzen, H.-M., Strähle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R. & Schütz, G. (1987) *Cell* **49,** 29–38.

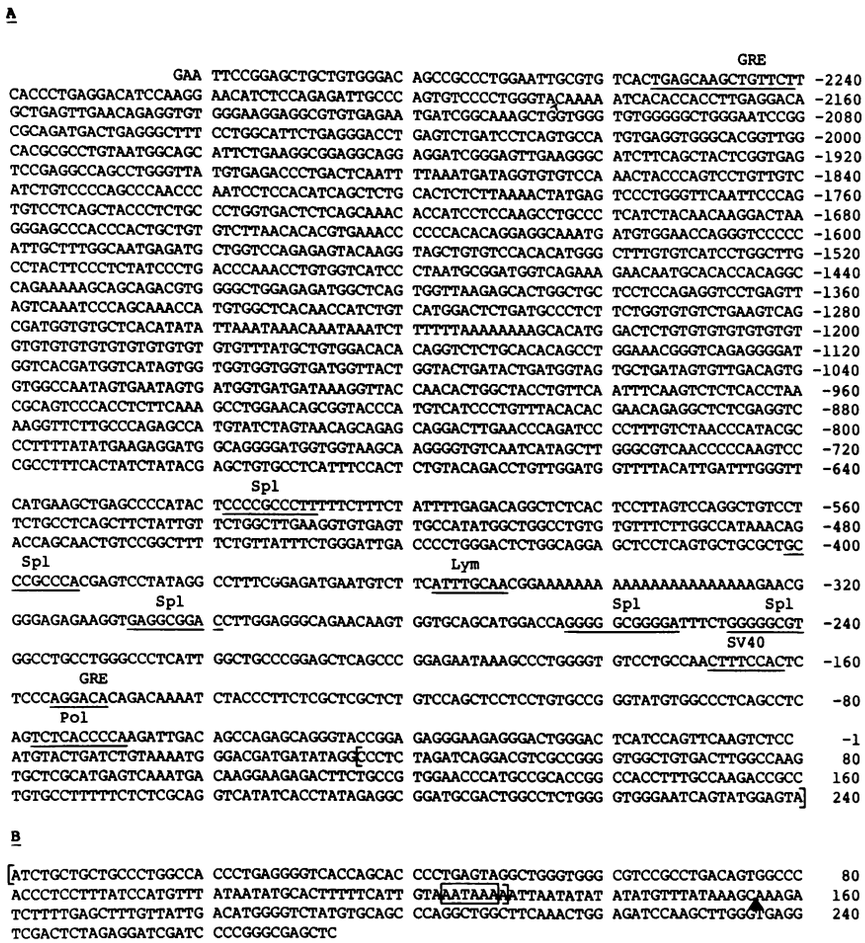

FIG. 4. Nucleotide sequences of 5′ (*A*) and 3′ (*B*) flanking regions of the serine dehydratase gene. The sequences were determined on both strands of the restricted fragments and fully overlapped. (*A*) The middle portion of the 3-kbp fragment of serine dehydratase genomic DNA was sequenced. The site of transcription initiation determined by S1 nuclease mapping is numbered +1. The sequence in brackets is the sequence identical with that of the most 5′ region of SDH2 (cf. Fig. 2). This sequence is followed by the first intron. (*B*) The *Eco*RV–*Eco*RI fragment located at the most 3′ end of serine dehydratase genomic clone pGEM 6.6 was sequenced and the *Eco*RV cutting site is numbered 1. The sequence in brackets is identical with that of the 3′ region of SDH2 (Fig. 2). Sequences identical with or similar to consensus sequences are underlined. Lym, lymphoid-specific promoter (28); Sp1, binding site for transcription factor Sp1 (29); SV40, enhancer core of 72-bp repeat sequence of simian virus 40 (30); GRE, glucocorticoid-responsive element (31); Pol, enhancer core of polyoma virus (32). A poly(A) addition signal is boxed, and the probable poly(A) addition site is indicated by an arrowhead.

30. Nordheim, A. & Rich, A. (1983) *Nature (London)* **303,** 674–679.
31. Kruys, V., Wathelet, M., Poupart, P., Contreras, R., Fiers, W., Content, J. & Huez, G. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6030–6034.
32. Karin, M., Haslinger, A., Holtgreve, H., Richards, R. I., Krauter, P., Westphal, H. M. & Beato, M. (1984) *Nature (London)* **308,** 513–519.
33. Berk, A. J. & Sharp, P. A. (1977) *Cell* **12,** 721–732.
34. Kadonaga, J. T., Jones, K. A. & Tjian, R. (1986) *Trends Biochem. Sci.* **11,** 20–30.
35. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
36. Takiguchi, M., Murakami, T., Miura, S. & Mori, M. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6136–6140.
37. Weiher, H., Konig, M. & Gruss, P. (1983) *Science* **219,** 626–631.
38. Prochownik, E. V. (1985) *Nature (London)* **316,** 845–848.
39. Wirth, T., Staudt, L. & Baltimore, D. (1987) *Nature (London)* **329,** 174–177.
40. Berget, S. M. (1984) *Nature (London)* **309,** 179–181.
41. Bergman, I. E. & Brawerman, G. (1977) *Biochemistry* **16,** 259–264.
42. Brawerman, G. (1987) *Cell* **48,** 5–6.
43. Huez, G., Bruck, C. & Cleuter, Y. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 908–911.
44. Shaw, G. & Kamen, R. (1986) *Cell* **46,** 659–667.
45. Pitot, H. C., Peraino, C., Lamar, C., Jr., & Kennan, A. L. (1965) *Proc. Natl. Acad. Sci. USA* **54,** 845–851.