Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 4369–4373, July 1987
Biochemistry

Molecular cloning and sequence of a cDNA coding for bovine lipoprotein lipase

MOTOTAKA SENDA*, KAZUHIRO OKA†, W. VIRGIL BROWN†, PRADMAN K. QASBA‡, AND YASUHIRO FURUICHI*§

*Department of Molecular Genetics, Nippon Roche Research Center, Kamakura, 247 Japan; †Department of Medicine, Mount Sinai School of Medicine, New York, NY 10029; and ‡Division of Cancer Biology and Diagnosis, National Cancer Institute, Rockville, MD 20851

Communicated by John J. Burns, March 9, 1987

ABSTRACT **Lipoprotein lipase (LPL; triacylglycero-protein acylhydrolase, EC 3.1.1.34) was purified from bovine milk. Synthetic oligonucleotides were prepared, based on the amino acid sequences of three peptides obtained from partial digestion of purified LPL, and were used as probes to isolate cDNA clones for LPL mRNA from a bovine mammary gland. One of the clones, pLPL-49R2, contains an insert cDNA (49R2) of about 3.2 kilobases (kb) that hybridizes to all three probes and encodes a polypeptide that includes the NH_2-terminal sequence of bovine LPL reported recently [Ben-Avram, C. M., Ben-Zeev, O., Lee, T. D., Hagga, K., Shively, J. E., Goers, J., Pedersen, M. E., Reeve, J. R. & Schotz, M. C. (1986) *Proc. Natl. Acad. Sci. USA* 83, 4185–4189]. Complete nucleotide sequence analysis revealed that cDNA insert 49R2 contains the entire coding region for LPL as well as a 3′ untranslated region of about 1.6 kb. The predicted amino acid sequence indicates that bovine LPL is a hydrophilic protein consisting of 450 amino acids (M_r 50,548) in its unglycosylated form. Blot hybridization analysis of poly(A)$^+$ mRNA from bovine mammary gland demonstrated that there are at least three sizes of LPL mRNAs—3.2, 2.5, and 1.7 kb—with the 2.5-kb mRNA being the most abundant. Restriction endonuclease mapping of other cDNA clones suggested that the variation in mRNA size results from differential utilization of polyadenylylation signals during mRNA processing.**

Lipoprotein lipase (LPL; triacylglycero-protein acylhydrolase, EC 3.1.1.34) is a lipolytic enzyme involved in the metabolism of triacylglycerol-rich lipoprotein particles (1). It hydrolyzes the triacylglycerol cores of chyromicrons and low density lipoproteins, generating free fatty acids available for extrahepatic tissues. LPL is a glycosylated protein of apparent M_r 55,000 that is synthesized and secreted from parenchymal cells of adipose tissue, heart, skeletal muscle, and mammary gland and is anchored on the luminal surface of the capillary endothelium by membrane-bound heparan sulfate (2, 3). The activity of LPL is markedly enhanced by apolipoprotein C-II present on the surface of triacylglycerol-rich lipoproteins (4). Recent findings (5) indicate that LPL is inhibited by tumor necrosis factor (cachectin). Detailed genetic analyses of several mutant mouse strains suggest that LPL is involved in the cause of atherosclerosis (6). LPL is also present in milk, but the reason for its presence is unknown (1).

Since the first description of canine LPL in 1943 (7), a sustained effort has been made to characterize its enzymatic activity and functions (for review, see refs. 1, 2, and 6). However, relatively little is known about the structure of LPL or the gene encoding LPL, or about the factors that govern its expression, despite clinical evidence that familial hypertriglyceridemia is a genetic disorder (8). Recently, Ben-Avram *et al.* (9) described the NH_2-terminal and partial amino acid sequence of bovine LPL and its homology to pancreatic lipase (10). We report here the isolation and characterization of a cDNA clone that contains the entire coding region of mature bovine LPL.

MATERIALS AND METHODS

Characterization of Bovine Milk LPL. *Purification.* LPL was purified from low-fat bovine skim milk according to the method described by Bengtsson and Olivecrona (11), with modifications. In brief, a 5000-fold purification of LPL was obtained by heparin-agarose column chromatography followed by phenyl-agarose column chromatography. The assay for LPL enzyme activity was carried out as described (12). Polyacrylamide gel electrophoresis according to Laemmli (13) was carried out in a slab gel, and the gels were stained with Coomassie brilliant blue R-250.

Digestion of LPL with trypsin and cyanogen bromide, and separation of peptides by reverse-phase HPLC. Purified LPL (50 μg) was treated with trypsin (0.5 μg) in 20 mM Tris/HCl buffer (pH 7) for 14 hr at 37°C. For chemical digestion, S-carboxymethylated LPL (50 μg) was incubated with cyanogen bromide (≈50 mg) in 50 μl of 70% formic acid 24 hr at room temperature. The digests were fractionated on a reverse-phase column (8.0 mm × 75 mm, Chemcosorbs ODS-H, Chemco Scientific, Osaka, Japan), using a 40-min linear gradient of 0–60% acetonitrile in 0.1% trifluoroacetic acid at a flow rate of 1 ml/min. The fractionated peptides were recovered by evaporation of the solvent.

Microsequence analysis. A gas-phase peptide/protein microsequencer [Applied Biosystems (Foster City, CA), model 470A] was used for automated Edman chemical degradations. The phenylthiohydantoin amino acid derivatives were identified by reverse-phase HPLC on an Aquasil SEQ-4K column (4.6 mm × 30 cm; Senshu, Tokyo) and were quantitated by an integration program on a Spectra Physics (San Jose, Ca) 4270 spectrophotometer.

Molecular Cloning of cDNA Encoding LPL. *cDNA library.* A cDNA library was prepared from poly(A)$^+$ mRNA obtained from lactating bovine mammary gland, using the Okayama–Berg procedure (14). Double-stranded cDNAs were ligated to plasmid pCD-X (15) and used to transform *Escherichia coli* MC1061 cells.

Oligonucleotide probes. Based on the amino acid sequence data, several oligodeoxyribonucleotide probes (21- to 23-mers) were synthesized using an Applied Biosystems 381A DNA synthesizer. These probe DNAs were 5′-^{32}P-labeled (≈10^8 cpm/μg) by polynucleotide kinase and were used in colony hybridization (16) to screen for bacterial clones that contain LPL-specific sequences. Hybridization was carried out overnight at 50°C in 6× SSC containing 5× Denhardt's

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.

Abbreviation: LPL, lipoprotein lipase.
§To whom reprint requests should be addressed.

Table 1. Purification of bovine milk LPL

Step	Volume, ml	Protein, mg	Activity, μmol/hr	Specific activity, μmol/hr per mg of protein
Skim milk	1800	25,090	52,200	2.1
Heparin-agarose column chromatography	275	358	36,300	10,139.7
Phenyl-agarose column chromatography	19	0.49	5,472	11,167.3

LPL activity was assayed by the method of Baginsky and Brown (12) and expressed as μmol of fatty acid released during a 60-min incubation at 37°C.

solution, 0.5% $NaDodSO_4$, and 20 μg of salmon sperm DNA per ml. (SSC is standard saline citrate; 1× SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.0. Denhardt's solution (1×) is 0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% bovine serum albumin.) Filters were washed in 6× SSC at 50°C for 2 hr with three changes of washing buffer.

DNA and RNA blot analyses. To verify that cDNA clones contained LPL-specific sequences, transblotting analysis by the Southern procedure (17) was carried out for individual plasmid DNAs, using 5′-^{32}P-labeled synthetic DNA probes.

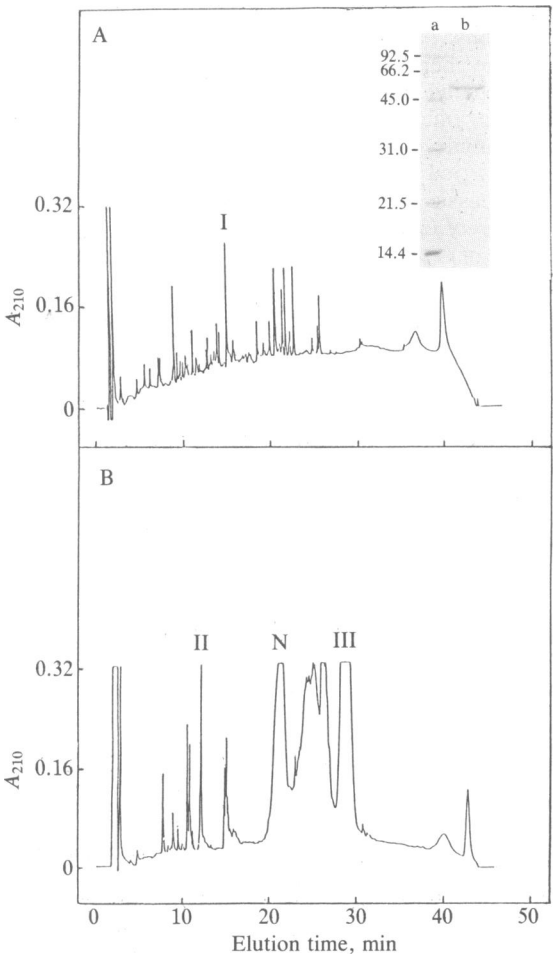

FIG. 1. Reverse-phase HPLC maps of peptides obtained by digestion of purified LPL (50 μg) with trypsin (*A*) or with cyanogen bromide. Peak fractions I, II, and III were further characterized by amino acid sequencing. Peak N contained the NH_2-terminal peptide of LPL. The amino acid sequences of these peaks are given in the text. (*Inset*) Gel electrophoresis of purified bovine milk LPL. Purified enzyme was electrophoresed in a 12% polyacrylamide gel in the presence of 0.1% $NaDodSO_4$ and 0.1% 2-mercaptoethanol. The gel was stained with Coomassie brilliant blue. Lane a: molecular weight markers ($M_r \times 10^{-3}$ at left). Lane b: ≈10 μg of LPL.

Similarly, RNA blot analysis (18) was performed for poly(A)$^+$ mRNA prepared from bovine mammary gland (19) to detect LPL-specific mRNA species. For this, the *Pst* I fragment [541 base pairs (bp)] from 49R2 cDNA (bases 313–853 in Fig. 3) was ^{32}P-labeled by nick-translation and used as the probe.

Nucleotide sequencing. Sequencing of the cloned LPL cDNA was carried out by the M13 procedure of Sanger (20).

RESULTS

Amino Acid Sequence Analysis of LPL. LPL was purified to homogeneity from bovine skim milk (Table 1; Fig. 1 *A Inset*). The purified LPL was then subjected to amino acid and peptide analyses as described in *Materials and Methods*. Our amino acid composition data were consistent with the data reported by Ben-Avram *et al.* (9). Also, our NH_2-terminal sequence data (a stretch of 22 amino acid residues: Asp-Arg-Ile-Thr-Gly-Gly-Lys-Asp-Phe-Arg-Asp-Ile-Glu-Ser-Lys-Phe-Ala-Leu-Arg-Thr-Pro-Glu) agreed with the 19 residues determined by Ben-Avram *et al.* and, as discussed in a later section, are also consistent with the sequence predicted from the nucleotide sequence of the cloned LPL cDNA. These results not only identify the purified polypeptide as LPL but also confirm that the NH_2 terminus of bovine milk LPL is not modified, for example, by N-acetylation.

To obtain more information about the amino acid sequence, we digested the purified protein with pancreatic trypsin or cyanogen bromide, and the resulting peptides obtained as discrete peaks by HPLC were sequenced.

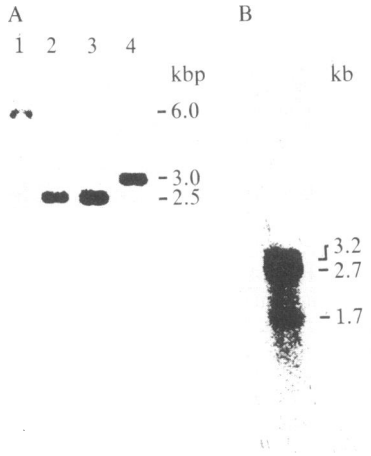

FIG. 2. Transfer blot analyses of cloned cDNA (*A*) and poly(A)$^+$ RNA (*B*) from bovine mammary gland. (*A*) cDNA clones pLPL-49R2 (lane 4), pLPL-49R3 (lane 2), and pLPL-54R1 (lane 3) were digested with restriction endonuclease *Bam*HI, electrophoresed in a 1% agarose gel, and hybridized with 5′-^{32}P-labeled probe A after transblotting to nitrocellulose paper (18). Lane 1 shows undigested pLPL-49R2. (*B*) Poly(A)$^+$ RNA was electrophoresed in a 1% agarose/6% formaldehyde gel and blotted onto a Hybond N filter (Amersham). The RNA was fixed on the filter by exposure to UV (300-nm) light for 5 min and hybridized with a ^{32}P-labeled *Pst* I DNA fragment (541-bp fragment from 5′-proximal region).

Peptides in peaks I, II, and III (Fig. 1) contained the sequences Ile-Thr-Gly-Leu-Asp-Pro-Ala-Gly-Pro-Asn-Phe--Glu-Tyr-Ala-Glu-, (Met)-Gly-Tyr-Glu-Ile-Asn-Lys-Val-Arg--Ala-Lys-Arg-Ser-Ser-Lys-Met-, and (Met)-Ala-Asp-Glu-Phe--Asn-Tyr-Pro-Leu-, respectively. Mixed oligodeoxyribonucleotide probes potentially coding for these peptides were prepared, based on codon degeneracy and frequency of eukaryotic codon usage (21): probe B (a 20-mer corresponding to a peptide in peak I) was a 48-fold-degenerate mixture of the sequence 5′-ATCACRGGSCTSGAYCCVGC-3′; probe C (a 21-mer corresponding to a peptide in peak II) was a 16-fold-degenerate mixture of the sequence 5′-ATGGGY-TAYGARATYAACAAG-3′; and probe A (a 23-mer corresponding to a peptide in peak III) was a 128-fold-degenerate mixture of the sequence 5′-ATGGCNGAYGARTTYAART-AYCC-3′. (Y is C or T; R is A or G; S is C or G; V is A, C, or G; and N is A, C, G, or T.)

Isolation and Characterization of cDNA Clones. About 600,000 transformants, derived from a cDNA library constructed from lactating bovine mammary gland poly(A)$^+$ RNA in Okayama–Berg vectors, were screened by using labeled oligonucleotides as hybridization probes. ^{32}P-labeled probe A was used as the primary probe to screen all transformants, and probes B and C were used as second and third sorting probes to select LPL-specific cDNA clones. Three cDNA clones were identified that reacted with probes A, B, and C, and one clone that reacted only with probes B and C. The cDNA inserts of these plasmids were excised by digestion with endonuclease *Bam*HI. Their sizes ranged from 1.7 to 3.2 kilobases (kb). The plasmid with the longest insert (3.2 kb) was designated pLPL-49R2; the two with medium-sized inserts (2.5 kb) were called pLPL-49R3 and pLPL-54R1. Upon Southern blot analysis, labeled probe A reacted with all inserts examined (Fig. 2*A*). In anticipation that cDNA insert 49R2 might contain the entire coding region for LPL, we analyzed this clone by restriction enzyme mapping and

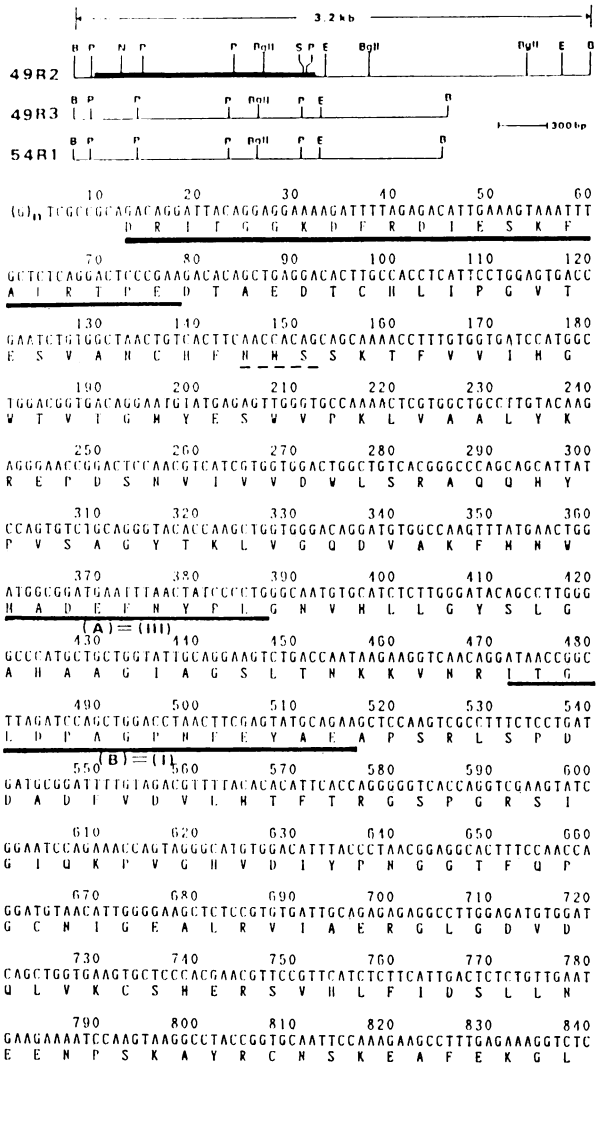

FIG. 3. Restriction endonuclease maps of the cDNA inserts 49R2, 49R3, and 54R1 are shown at the top of the left-hand column: P, S, BglI, N, B, and E indicate *Pst* I, *Sma* I, *Bgl* II, *Nco* I, *Bam*HI, and *Eco*RI restriction sites, respectively; the bold line in 49R2 indicates the region encoding the LPL amino acid sequence. The remainder of the figure shows the complete nucleotide sequence of the 49R2 cDNA and the predicted amino acid sequence of the bovine LPL (standard one-letter amino acid symbols, given below the nucleotide sequence). The sequence of the NH_2-terminal peptide (see Fig. 1*B*, peak N) is underlined. Also underlined are the regions corresponding to the peptides (I, II, and III) on which the synthetic hybridization probes A (23-mer), B (21-mer), and C (20-mer) were based. Broken underlining shows the two potential N-glycosylation sites. The two potential cleavage/polyadenylylation signals (AATAAA) at bases 2033–2038 and 2882–2887 are underlined. The nucleotide sequence was determined by reading both strands except for the region (bases 2111–2342) in parentheses.

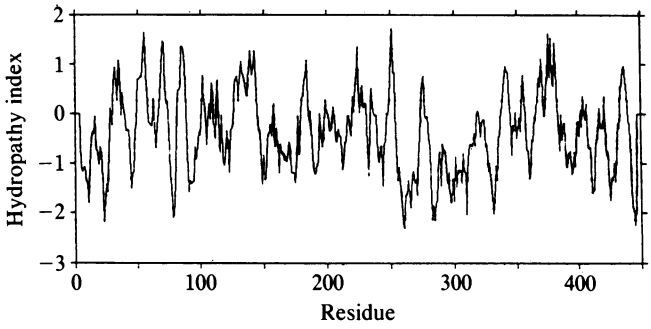

FIG. 4. Hydrophobic and hydrophilic regions of bovine milk LPL. The predicted primary amino acid sequence of LPL was computer-scanned for hydrophobicity, using the program of Kyte and Doolittle (23). The search length was nine amino acids. Hydrophobicity (hydropathy) increases with distance above the abscissa.

nucleotide sequencing. Restriction mapping of the other two inserts (49R3 and 54R1) showed that they lacked the 3′-terminal regions of 49R2 (Fig. 3).

Nucleotide Sequence of Cloned cDNA. The major emphasis in sequence analysis was to locate the coding region for the NH_2-terminal and three internal peptides of LPL whose amino acid sequence had been established by microsequencing. As shown in Fig. 3, the complete nucleotide sequence of 49R2 (2903 bases) was determined. The region coding for the NH_2-terminal peptide Asp-Arg-Ile-Thr-Gly-Gly-Lys-Asp-Phe-Arg-Asp-Ile-Glu-Ser-Lys-Phe-Ala-Leu-Arg-Thr-Pro-Glu was found at the 5′ end of the cDNA insert (bases 13–78). The upstream region that should code for the signal peptide was absent in this clone, except for the sequence TCGC-CGCA that encodes three amino acids (Xaa-Ala-Ala, where Xaa is Phe, Leu, Ile, or Val) absent from the mature LPL. Regions that code for peptides I, II, and III (Fig. 1) were found at bases 472–516, 874–921, and 361–387 (Fig. 3), respectively. All four peptides were in the same open reading frame, which was interrupted by a terminator TGA at bases 1363–1365. There were many stop codons thereafter. This reading frame encompasses 1350 bases capable of encoding a polypeptide consisting of 450 amino acid residues. These sequence data predict that the mature bovine LPL has a molecular weight of 50,548, in its unglycosylated form.

The *Bgl* II–*Bam*HI fragment located at the 3′ end of 49R2 cDNA contains a poly(A) sequence and an AATAAA sequence 15 bases upstream of the poly(A), characteristics of eukaryotic mRNA. There is another AATAAA sequence in the middle of the 3′ noncoding region (residues 2033–2038), which is preceded by several AATAAA analogue sequences and followed by a "G/T cluster," indicative of a strong mRNA cleavage/polyadenylylation site (22).

mRNA Analysis by Blot Hybridization. To determine the original size of LPL mRNA, total poly(A)$^+$ RNA extracted from lactating bovine mammary glands was analyzed by blot hybridization (Fig. 2*B*) using as probe a ^{32}P-labeled *Pst* I–*Pst* I fragment (541 base pairs) present in all cloned cDNAs. The major hybridizable RNA band migrated at a position corresponding to 2.5 kb. Two other minor, but distinct, RNA bands of 3.2 and 1.7 kb were also detected (Fig. 2*B*).

Judging from their size and restriction maps, cDNA 49R2, and the two medium-sized cDNAs, 49R3 and 54R1, probably derive from the 3.2-kb and 2.5-kb mRNAs, respectively. While these three cDNAs share the same coding region (from restriction maps) and all contain poly(A) at the 3′ ends (data not shown), both 49R3 and 59R1 cDNA have a shorter 3′ noncoding region.

Structural Features of Bovine LPL. LPL is known to be a glycosylated protein associated with the luminal surface of capillary endothelium via membrane-bound heparan sulfate (1–3). The predicted amino acid sequence of bovine LPL contains two potential sites [Asn-Xaa-(Thr/Ser)] for N-linked glycosylation at amino acid residues 45–47 and 361–363. Hydropathicity plots (Fig. 4), constructed according to Kyte and Doolittle (23), indicate that bovine LPL is very hydrophilic, with the two potential glycosylation sites occurring at β-turns in hydrophilic domains. There are two regions rich in basic amino acids in the COOH-terminal region of LPL (amino acid residues 262–308 and 405–439) that could associate with the acidic domain of heparan sulfate. Both of these regions have an α-helix configuration [based on analysis as described by Chou and Fasman (24)].

DISCUSSION

In this paper we describe the cloning of mRNAs that encode bovine LPL. One of the cloned cDNAs (49R2) apparently contains the complete nucleotide sequence coding for mature LPL, although it lacks the 5′ region encoding the signal peptide required for translocation into the lumen of the rough endoplasmic reticulum. Curiously, the 49R2 cDNA was found to contain a very long 3′ untranslated region. Two other cDNA clones were apparently identical to 49R2 with respect to the LPL coding region but contained shorter 3′ untranslated regions that include terminal poly(A). These results imply that there are multiple species of LPL mRNA in bovine mammary gland. In fact, blot analysis of poly(A)$^+$ RNA from bovine mammary gland showed that there are at least three sizes of LPL mRNAs. Judging from the intensity of the hybridized bands (Fig. 2*B*), the medium-sized mRNA is the most abundant species in bovine mammary gland, whereas cDNA insert 49R2 seems to correspond to the longest (3.2-kb) mRNA species.

The biological significance and biogenesis of these three mRNA species remain to be elucidated. However, these mRNAs most likely derive from differential transcription termination, as shown in the adenovirus major late transcription unit (25), since the cDNA clones obtained in this study appear to be different only in the 3′ noncoding region. Partial sequence analysis (data not shown) of the medium-size cDNA, 49R3, shows that the upstream AATAAA (at bases 2033–2038) apparently has been used for the cleavage/polyadenylylation of the template mRNA. A similar observation concerning various sizes of mRNA from bovine mammary gland was reported (26) for mRNAs encoding *N*-acetylglucosamine (β1-4)galactosyltransferase.

Recently, Reddy *et al.* (27) reported the amino acid sequence of a peptide that contains a serine residue reactive to diisopropylphosphorofluoridate, a potential active site of bovine LPL. However, we did not locate the corresponding peptide in our predicted LPL sequence. In fact, no such sequence could be found even in the other two reading frames of the cloned LPL cDNA 49R2. Further analyses of genomic clones and assignment of individual cDNA clones to LPL polypeptides should open the way to understanding of regulation of LPL gene expression and clinical application for either diagnostics or as a remedy for genetic disorders.

We are grateful to Dr. J. Paterniti for encouragement, Ms. L. A. Houghton for technical assistance, and Ms. M. Shida for preparation of the manuscript. This work was supported in part by National Institutes of Health Grant HL32144.

1. Nilsson-Ehle, P., Garfinkel, A. S. & Schotz, M. C. (1980) *Annu. Rev. Biochem.* **49**, 667–693.
2. Cryer, A. (1981) *Int. J. Biochem.* **13**, 525–541.
3. Chimada, K., Gill, P. J., Silbert, J. E., Douglas, W. H. & Fanburg, B. L. (1981) *J. Clin. Invest.* **68**, 995–1002.
4. La Rosa, J. C., Levy, R. I., Herbert, P., Lux, S. E. & Fredrickson, D. S. (1970) *Biochem. Biophys. Res. Commun.* **41**, 57–62.

5. Mahoney, J. R., Jr., Beutler, B. A., LeJrang, N., Vine, W., Ikeda, Y., Kawakami, M. & Cerami, A. (1985) *J. Immunol.* **134,** 1673–1675.
6. Oka, K. & Paterniti, J. R., Jr. (1986) *Drug Dev. Res.* **7,** 377–392.
7. Hahn, P. F. (1943) *Science* **98,** 19–21.
8. Beaumont, J.-L., Carlson, L. A., Cooper, G. R., Feifar, Z., Fredrickson, D. S. & Strasser, T. (1970) *Bull. W.H.O.* **43,** 891–915.
9. Ben-Avram, C. M., Ben-Zeev, O., Lee, T. D., Haaga, K., Shively, J. E., Goers, J., Pedersen, M. E., Reeve, J. R. & Schotz, M. C. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4185–4189.
10. De Caro, J., Boudouard, M., Guidoni, A., Desnuelle, P. & Rovery, M. (1981) *Biochim. Biophys. Acta* **660,** 148–150.
11. Bengtsson, G. & Olivecrona, T. (1977) *Biochem. J.* **167,** 109–119.
12. Baginsky, M. L. & Brown, W. V. (1979) *J. Lipid Res.* **20,** 548–556.
13. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
14. Okayama, H. & Berg, P. (1982) *Mol. Cell. Biol.* **2,** 161–170.
15. Okayama, H. & Berg, P. (1983) *Mol. Cell. Biol.* **3,** 280–289.
16. Grunstein, M. & Hogness, D. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3961–3965.
17. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
18. Alvine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5350–5354.
19. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
20. Sanger, F. (1981) *Science* **214,** 1205–1210.
21. Grantham, R., Gautier, C., Gouy, M., Jacobzome, M. & Mercier, R. (1981) *Nucleic Acids Res.* **9,** r43–r74.
22. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) *Cell* **41,** 349–359.
23. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
24. Chou, P. Y. & Fasman, G. C. (1978) *Adv. Enzymol.* **47,** 45–148.
25. Nevins, J. R. & Darnell, J. E., Jr. (1978) *Cell* **15,** 1477–1493.
26. Narimatsu, H., Sinha, S., Brew, K., Okayama, H. & Qasba, P. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4720–4724.
27. Reddy, M. N., Maraganore, J. M., Meredith, S. C., Heinrikson, R. L. & Kezdy, F. J. (1986) *J. Biol. Chem.* **261,** 9678–9683.