Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 1526–1530, March 1987
Biochemistry

Cloning of rat hepatic lipase cDNA: Evidence for a lipase gene family

(lipoprotein lipase/pancreatic lipase)

MICHAEL C. KOMAROMY AND MICHAEL C. SCHOTZ*

Research, Veterans Administration, Wadsworth Medical Center, Los Angeles, CA 90073; and Department of Medicine, University of California, Los Angeles, CA 90024

Communicated by Richard E. Dickerson, November 20, 1986

ABSTRACT **Clones for rat hepatic lipase were isolated by probing a rat liver cDNA library in λgt11 with an oligonucleotide synthesized on the basis of a partial peptide sequence. The cloned messenger codes for a protein of 472 amino acids plus a hydrophobic leader sequence of 22 amino acids. The unglycosylated protein has a predicted molecular weight of 53,222 and contains two potential sites for N-glycosylation. The protein bears striking regions of homology with other known lipases and contains peptide sequences that have been implicated in lipid binding. The homologous mRNA is present in liver tissue but no detectable mRNA is observed in the adrenal gland, despite the reported presence of hepatic lipase in both the liver and the adrenal gland. No mRNA was seen in any of a variety of other tissues.**

Hepatic lipase, localized primarily on the sinusoidal surfaces of the liver, functions in the metabolism of circulating lipoproteins (1). Hepatic lipase activity has also been detected in several extrahepatic tissues, including adrenal gland and ovary (2–4). The presumed function of the enzyme is the hydrolysis of triglycerides in intermediate density lipoproteins (IDL) and of phospholipids in high density lipoproteins (HDL_2) (5). The capacity of purified hepatic lipase to catalyze hydrolysis of phospholipids and mono-, di-, and triglycerides (6, 7) is consistent with the putative function. The action of hepatic lipase on HDL_2 is presumed to result in the production of HDL_3. Also, hepatic lipase may participate in the clearance of circulating very low-density lipoproteins (8) and chylomicron remnants (9).

Hepatic lipase from rat liver has been purified to homogeneity (7, 10). The enzyme has an apparent molecular weight of 53,000 (7, 11) and contains approximately 8% carbohydrate (12). Its activity is characterized by salt resistance, an alkaline pH maximum, and binding to heparin. Hepatic lipase has been shown to have physical and biochemical properties different from lipoprotein lipase, an extrahepatic lipase that catalyzes the initial hydrolysis of circulating triglyceride-rich lipoproteins. However, some partial amino acid sequences of hepatic lipase are identical to those of both lipoprotein lipase and pancreatic lipase (11), an enzyme involved in the hydrolysis of triglycerides in the intestine. It seems likely that these highly conserved regions are critical for the function of these lipases.

Using an oligonucleotide probe constructed from a partial amino acid sequence of hepatic lipase, we screened a rat liver complementary DNA (cDNA) library. Hepatic lipase cDNA clones were selected and sequenced, and the complete amino acid sequence was obtained.

MATERIALS AND METHODS

Materials. *Escherichia coli* Y1088 and Y1090 and an oligo(dT)-primed cDNA library from rat liver poly(A)$^+$ RNA, prepared in the λgt11 vector system, were provided by Camilla Heinzman, Todd Kirchgessner, and Richard Tanaka (University of California, Los Angeles). Radioisotopes and restriction enzymes were obtained from Amersham and Pharmacia, respectively.

cDNA Screening and Analysis. Partial amino acid sequences were determined from the digestion of purified hepatic lipase with subtilisin (11). A single 42-base oligonucleotide probe was constructed on the basis of a 14 amino acid sequence (Fig. 1) and codon usage tables for mammalian genes (13). In addition to codon usage assumption, G·T base pairing was allowed in designing the probe. A total of 9×10^5 plaques were screened with the end-labeled 42-base oligonucleotide probe. Plaque replicas on nitrocellulose filters were made by the method of Benton and Davis (14). The filters were baked at 80°C for 2 hr and incubated for 3 hr at 37°C in 30% (vol/vol) formamide/900 mM NaCl/90 mM trisodium citrate/5× concentrated Denhardt's solution/50 mM sodium phosphate, pH 6.5, containing salmon sperm DNA at 100 μg/ml, in a total volume of 200 ml. The ^{32}P-labeled probe (2×10^8 cpm/μg) was hybridized to the filters for 17 hr at 37°C in 200 ml of the same solution. The filters were washed twice at room temperature in 500 ml of 150 mM NaCl/15 mM trisodium citrate/0.1% $NaDodSO_4$ for 30 min per wash, and then six times at 37°C in 300 ml of the same solution for 30 min per wash. The filters were exposed to Du Pont Cronex 4 film with Du Pont Hi-Plus intensifying screens at −70°C for 20 hr. DNA from clones that had hybridized with the probe was isolated from plate lysates and subcloned in phage M13 mp19 for sequence analysis by dideoxy chain termination (15). Further subcloning was done in the same vector as needed to get the complete sequence of both strands.

RNA Isolation, Electrophoresis, and Blot Hybridization. RNA was prepared from rat tissues by homogenization in guanidinium thiocyanate followed by centrifugation over a cesium chloride cushion (16); poly(A)$^+$ RNA was enriched by oligo(dT)-cellulose chromatography. RNA was electrophoresed in 1.0% agarose gels in the presence of formaldehyde (17) and transferred to either nylon or nitrocellulose membranes by standard blotting procedures. Hybridizations with the ^{32}P-labeled cDNA probe were carried out for 20 hr at 42°C in 50% formamide/10% dextran sulfate/300 mM NaCl/30 mM trisodium citrate/0.1% $NaDodSO_4$/1 mM EDTA containing salmon sperm DNA at 100 μg/ml. A series of 10-min washes was performed at 68°C in 300 mM NaCl/30 mM trisodium citrate, 150 mM NaCl/15 mM trisodium citrate, 75

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 at: Veterans Administration, Wadsworth Medical Center, Lipid Research Laboratory, Bldg. 113, Rm. 312, Los Angeles, CA 90073.

```
Peptide Sequence      Leu Trp Asp Thr Glu Pro His Tyr Ala Gly Leu Ile Leu Lys
Coding Sequence   5'-CTA TGG GAC ACC GAA CCC CAC TAC GCC GGC CTA ATC CTA AAA
Complement(probe) 3'-GAT ACC CTG TGG CTT GGG GTG ATG CGG CCG GAT TAG GAT TTT
```

FIG. 1. Peptide sequence, derived coding sequence, and actual probe sequence. The probe is the complement of the derived coding sequence, so that it could be used to screen both plaques and RNA blots. The derived coding sequence was based on codon usage tables for mammalian genes and allowing G·T base pairing. Comparison with the actual coding sequence (Fig. 3) shows that the probe is an 81% match (36/42 bases).

mM NaCl/7.5 mM trisodium citrate, 30 mM NaCl/3 mM trisodium citrate, and 15 mM NaCl/1.5 mM trisodium citrate. Autoradiography was done as described above.

RESULTS

The 42-base oligonucleotide probe (Fig. 1), end-labeled with ^{32}P, was used to isolate 12 clones from an adult rat liver cDNA library in λgt11. The inserts fell into three size classes. The five largest isolates were 1.65 kilobases (kb) in length and contained identical 5′ untranslated regions, suggesting that either these clones represent the complete 5′ end of the mRNA or they are all amplified isolates of a single original clone that may not contain the complete 5′ untranslated region. Primer extension analysis of the mRNA should resolve this question. Fig. 2 shows the partial restriction map and sequencing strategy for one of these cloned DNAs, C141.10. Also shown in Fig. 2 are the smaller clones C141.1 and C141.7, which appear by complete sequencing to be entirely contained in C141.10. Clone C141.1 and three others formed a size class of 1.1-kb inserts, and clone C141.7 and two others formed a 450- to 550-base size class.

The nucleotide sequence of C141.10 and the derived amino acid sequence are shown in Fig. 3. Protein sequences of the amino-terminal portion and of four internal peptides generated by proteolytic digestion of hepatic lipase were compared with the predicted amino acid sequence of this putative cDNA clone for hepatic lipase. All five sequences, consisting of 87 unambiguously identified amino acids (11), were present, intact, in regions distributed throughout the translated protein (see Fig. 3). This finding verifies the identity of the clone as rat hepatic lipase. The cDNA clone contains a 15-base 5′ untranslated region followed by the initiator codon for methionine. This is followed by a relatively hydrophobic stretch of 22 amino acids. The amino terminus of the mature protein, as identified by protein sequencing (11), begins at the glycine residue after the presumptive leader sequence. The protein encoded by the major open reading frame contains 472 amino acids with a calculated molecular weight of 53,222. This is followed by a 3′ untranslated region of 142 bases, including an AATAAA polyadenylylation signal (18) located 21 bases from the poly(A). Translation in other reading frames was interrupted by several stop codons. All 12 isolates contained a poly(A) tail at the same position, following nucleotide 1639.

The predicted sequence shows two potential N-glycosylation sites (Fig. 3) (Asn-Xaa-Ser, Asn-Xaa-Thr) (19), but it is not known if both are glycosylated in the native protein. Rat hepatic lipase migrates on $NaDodSO_4$/polyacrylamide gels with an apparent molecular weight of 53,000 (11). The estimated carbohydrate content is 8% (12). Thus, if fully glycosylated, the mature protein should have a molecular weight of about 57,500. The apparent molecular weight of 53,000 may reflect the anomalous migration of glycosylated proteins on $NaDodSO_4$/PAGE (20).

Blot hybridization analysis of both liver poly(A)$^+$ RNA and liver total RNA reveals a single hybridizing band at about 1750 nucleotides (Fig. 4). Assuming a typical mRNA poly(A) content of about 100 nucleotides, 1650 nucleotides would make up the hepatic lipase coding and untranslated regions. The largest clones reported here are 1639 nucleotides, suggesting that these isolates terminate at or near the actual 5′ end of the mRNA. The structural features of the translated protein reported above argue strongly that the entire protein coding region is included in the largest clones.

To determine whether hepatic lipase mRNA is present in extrahepatic tissue, blot hybridization analysis was performed. The adrenal gland contains approximately ¼ to ⅓ the hepatic lipase activity of liver tissue on a per gram basis (2); however, no detectable mRNA was found in the adrenal gland. The other tissue RNAs shown in Fig. 4 also show no detectable hybridization and are not known to contain the enzyme.

DISCUSSION

A full-length cDNA clone for rat hepatic lipase has been isolated and characterized. Although the actual 5′ end of the mRNA has not been mapped, the clone corresponds in size to the mRNA, and the five largest isolates contain the identical 5′ terminus. In addition, the clone contains a coding region for a potential hydrophobic leader sequence. Comparisons of the partial amino acid sequence of the mature protein with the predicted amino acid sequence of the clone verified the identity of the clone.

Comparison of the hepatic lipase amino acid sequence with the sequence of other lipases further substantiates the hypothesis (11) that rat hepatic lipase, bovine lipoprotein lipase, and porcine pancreatic lipase are members of a gene family. Fig. 5 shows a comparison of parts of the sequences of these

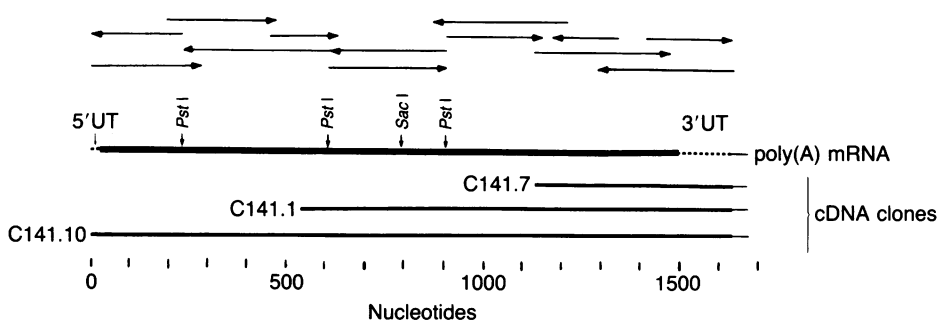

FIG. 2. Hepatic lipase mRNA, cDNA clones, and sequencing strategy. C141.7, C141.1, and C141.10 are cDNA clones representative of the three size classes (see text). The arrows denote the directions and approximate lengths of dideoxy sequencing. The partial restriction map shows the sites used for subcloning for sequence analysis. UT, untranslated region. In the cases in which arrows do not start at restriction sites, 17-mer oligonucleotide primers were constructed and used for the extensions.

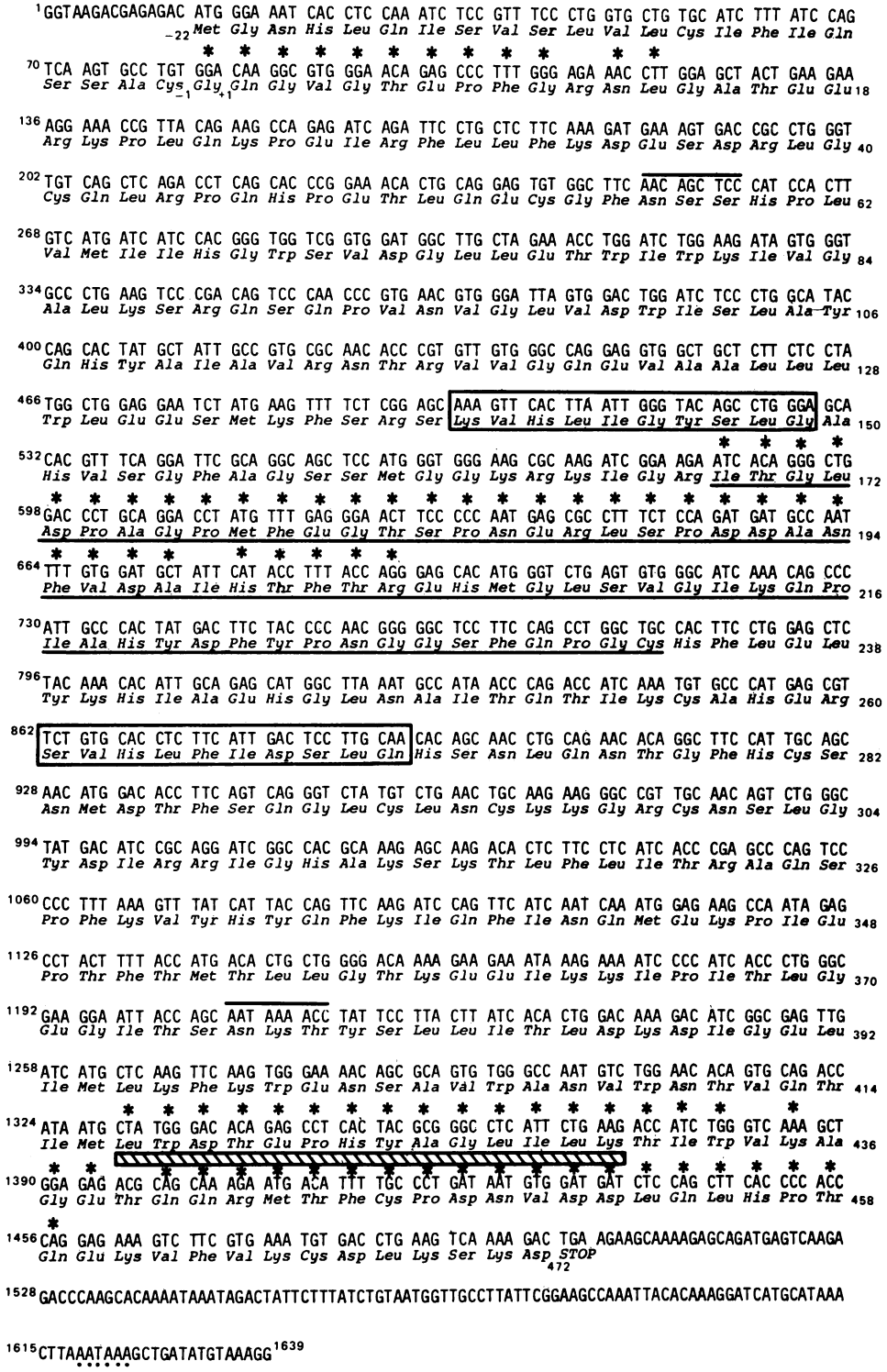

FIG. 3. Hepatic lipase nucleotide sequence and the predicted amino acid sequence. The superscript numbers refer to the nucleotide positions, and the subscript numbers refer to the amino acid positions. The sequence from −22 to −1 is the putative hydrophobic leader sequence. The probe sequence is underlined by a hatched box. The two potential N-glycosylation sites are overlined (amino acids 57–59 and 376–378). The polyadenylylation signal (AATAAA) is underdotted (nucleotides 1619–1624). The central homology region (see text) is underlined (amino acids 169–233). The lipid-binding regions are boxed (amino acids 140–149 and 261–270). Asterisks over the nucleotides mark amino acid sequences unambiguously confirmed by identity with direct peptide sequencing of the mature protein (11).

three proteins. In the regions shown the homology is high, with substantial stretches of identity. This striking homology, highly conserved in lipases of different functions and different species, suggests that these sequences are essential for catalytic activity. In addition to the homology shown in Fig. 5, a sequence in porcine pancreatic lipase, believed to form part of the interfacial lipid-binding region (21), is strikingly similar to sequences in other lipases—rat hepatic lipase, bovine lipoprotein lipase (11), human lecithin–cholesterol acyltransferase (19), and rat lingual lipase (22) (Fig. 6). The rat hepatic lipase sequence composed of residues 140–149 closely resembles porcine pancreatic lipase residues 145–154, while the rat hepatic lipase sequence from residues 261–270 is essentially identical to a sequence present in

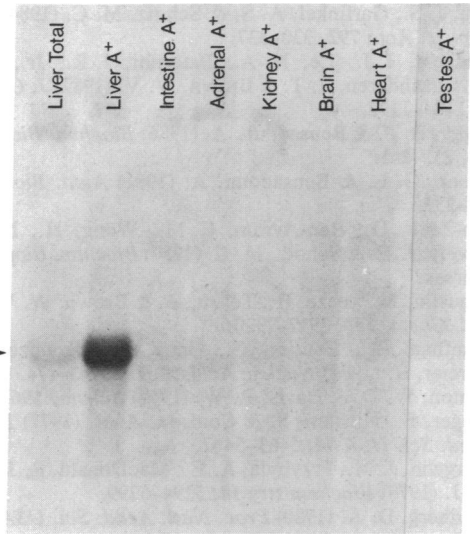

FIG. 4. Size and tissue distribution of hepatic lipase RNA. For each RNA, 10 μg was run, transferred, and hybridized. A^+ refers to poly(A)$^+$ RNA. Total RNA was not selected with oligo(dT). Markers (not shown) were *E. coli* 16S and 23S rRNA and yeast 18S and 28S RNA. The arrow at the left indicates the position of 18S rRNA. The blot was probed with the full-length cDNA clone.

bovine lipoprotein lipase. Except for the homologies shown in Figs. 5 and 6, computer dot-matrix analyses showed no other sequence homology between these lipases. However, the complete structure of lipoprotein lipase is not yet available. In addition, a self-comparison dot-matrix analysis of the hepatic lipase sequence shows no internal homologies other than the potential lipid-binding regions.

We have compared the hepatic lipase protein sequence with the sequences of other known heparin-binding proteins and have found no evidence for a specific heparin-binding region. For example, the sequence Ser-His-Leu-Arg-Lys-Leu-Arg-Lys-Arg, which forms the heparin-binding site of human apolipoprotein E3 (23), does not appear in the hepatic lipase sequence. There are no long stretches of positively charged amino acids, which may be expected to bind negatively charged heparin. The longest such region is three amino acids (amino acids 163–165). It is possible that the hepatic lipase heparin-binding site is formed by protein folding and is not apparent from the primary structure.

It has been noted that incubation of purified rat hepatic lipase with collagenase selectively decreased triglyceride lipase activity with no effect on the monoglyceride hydrolase

Rat lingual lipase	[163]Lys Ile His Tyr Val Gly His Ser Gln Gly
Human LCAT	[173]Pro Val Phe Leu Ile Gly His Ser Leu Gly
Porcine pancreatic lipase	[145]Asn Val His Val Ile Gly His Ser Leu Gly
Rat hepatic lipase 1	[140]Lys Val His Leu Ile Gly Tyr Ser Leu Gly
Rat hepatic lipase 2	[261]Ser Val His Leu Phe Ile Asp Ser Leu Gln
Bovine lipoprotein lipase	Ser Val His Leu Phe Ile Asp Ser Leu Leu

FIG. 6. Amino acid homologies between the lipid-binding region of porcine pancreatic lipase and other known lipase sequences. LCAT, lecithin–cholesterol acyltransferase. The numbers refer to amino acid positions; the lipoprotein lipase sequence is known from partial peptide sequence and is therefore not numbered. Hepatic lipase contains two homologous regions, which are designated 1 and 2.

activity (24). Since collagenase cleaved hepatic lipase into several fragments, it could be that hepatic lipase contains a collagen-like region that can serve as a substrate for collagenase. No collagenase recognition site is seen in the hepatic lipase sequence (25); perhaps a protease contaminant in the collagenase preparation accounts for cleavage of the hepatic lipase.

The amino acid sequence of the active site region of bovine lipoprotein lipase has been determined (26). The sequence contains a serine residue that can be modified with diisopropylphosphorofluoridate, an inactivator of serine esterases that also eliminates lipoprotein lipase activity against triacylglycerols. The lipoprotein lipase sequence contains the partial sequence Gly-Gly-Ser-Pro-Asn-Gln, in which the serine is the modified residue. Hepatic lipase contains the sequence Gly-Thr-Ser-Pro-Asn-Glu in the central homology region (amino acids 181–186). The sequence comparison in Fig. 5 shows that this sequence bears little homology with the aligned regions of pancreatic lipase and lipoprotein lipase; in fact, this sequence shows the least homology in this central region. On the basis of the alignment of the three enzymes in Fig. 5, it is clear that the active site as described by Reddy *et al.* (26) must be elsewhere in the lipoprotein lipase molecule. However, it is still possible that this region contains the active site for hepatic lipase. Site-directed mutagenesis of the hepatic lipase serine residue at position 183 should clarify this question.

The results of the blot hybridization analysis (Fig. 4) indicate that hepatic lipase is synthesized exclusively in the liver. Enzymatic evidence suggests that hepatic lipase activity is present in the adrenal gland in approximately the same amount as in liver on a weight basis (2, 3). From experiments with monoclonal antibodies to hepatic lipase it is concluded that the enzymes in liver and adrenal are identical (4). Although no mRNA was detected in the adrenal gland, it is possible that synthesis of the enzyme could occur at mRNA

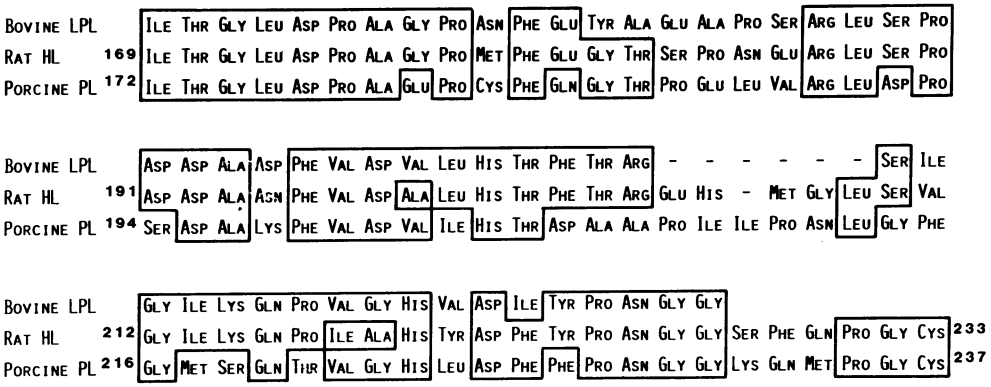

FIG. 5. Central amino acid homology regions of rat hepatic lipase (HL), bovine lipoprotein lipase (LPL), and porcine pancreatic lipase (PL). Identical residues are boxed. Numbers refer to amino acid positions; the lipoprotein lipase sequence is known from partial peptide sequencing and is therefore not numbered. The dashes were inserted to optimize the alignment and do not represent unidentified residues. Alignments were made by visual inspection. Overall identities are as follows: HL–LPL, 40/54 (74%); HL–PL, 34/65 (52%); PL–LPL, 30/54 (56%).

levels below our detection sensitivity. Alternatively, the adrenal enzyme activity may be due to a different lipase that is closely related to hepatic lipase structurally.

Hepatic lipase contains two internal homologous sequences, which also show considerable homology to the presumed lipid-binding region of pancreatic lipase (Fig. 6). These may represent binding sites for different lipid substrates. Alternatively, both proposed lipid-binding regions contained in hepatic lipase may be needed for a stable enzyme–substrate complex. The other lipases shown in Fig. 6 contain only a single copy of this proposed lipid-binding sequence. The other enzymes, with the exception of lingual lipase, require a protein cofactor for full activity. Hepatic lipase may be able to bind lipids more efficiently than other lipases and thus the protein cofactor requirement is eliminated. Mutagenesis of the amino acids in these regions, and in particular serine residues 147 and 268, should clarify this possibility and increase our understanding of the functional domains within the lipase gene family.

We are indebted to Bruce Kaplan and Peter Walker (Beckman Research Institute of the City of Hope, Department of Molecular Genetics) for synthesis of oligonucleotide probes. We are also grateful for the contribution of Chad Miller Ben-Avram, Osnat Ben-Zeev, and Howard Wong to the project. In addition, we are grateful to Diana Quon for preparing several tissue RNAs; Ruth West for preliminary screening of the liver library; and Todd Kirchgessner, Jake Lusis, and Ron Eng for their aid in designing the oligonucleotide probes. This study was supported, in part, by the National Institutes of Health (HL28481), the American Heart Association, Greater Los Angeles Affiliate (492IG14), and the Veterans Administration.

1. Jackson, R. L. (1983) in *The Enzymes*, ed. Boyer, P. (Academic, New York), Vol. 16, 141–181.
2. Jansen, H. & DeGreef, W. J. (1981) *Biochem. J.* **196,** 739–745.
3. Cordle, S. R., Yeaman, S. J. & Clegg, R. A. (1983) *Biochim. Biophys. Acta* **753,** 213–219.
4. Persoon, N. L. M., Sips, H. J., Hulsmann, W. C. & Jansen, H. (1986) *Biochim. Biophys. Acta* **875,** 286–292.
5. Kinnunen, P. K. J. (1984) in *Lipases*, eds. Borgström, B. & Brockman, H. L. (Elsevier, Amsterdam), pp. 307–328.
6. Laboda, H. M., Glick, J. M. & Phillips, M. C. (1986) *Biochim. Biophys. Acta* **876,** 232–242.
7. Twu, J.-S., Garfinkel, A. S. & Schotz, M. C. (1984) *Biochim. Biophys. Acta* **792,** 330–337.
8. Goldberg, I. J., Le, N.-A., Paterniti, J. R., Jr., Ginsberg, H. N., Lindgren, F. T. & Brown, W. V. (1982) *J. Clin. Invest.* **70,** 1184–1192.
9. Daggy, B. P. & Bensadoun, A. (1986) *Biochim. Biophys. Acta* **877,** 252–261.
10. Jensen, G. L. & Bensadoun, A. (1981) *Anal. Biochem.* **113,** 246–252.
11. Ben-Zeev, O., Ben-Avram, C. M., Wong, H., Nikazy, J., Shively, J. E. & Schotz, M. C. (1987) *Biochim. Biophys. Acta*, in press.
12. Augustin, J., Freeze, H., Tejada, P. & Brown, W. V. (1978) *J. Biol. Chem.* **253,** 2912–2920.
13. Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. & Mercier, R. (1981) *Nucleic Acids Res.* **9,** r43–r74.
14. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
16. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
17. Goldberg, D. A. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5794–5798.
18. Proudfoot, N. & Brownlee, G. (1981) *Nature (London)* **252,** 359–362.
19. McLean, J., Fielding, C., Drayna, D., Dieplinger, H., Baer, B., Kohr, W., Henzel, W. & Lawn, R. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2335–2339.
20. Schneider, W. J., Beisiegel, U., Goldstein, J. L. & Brown, M. S. (1982) *J. Biol. Chem.* **257,** 2664–2673.
21. Guidoni, A., Benkouka, F., DeCaro, J. & Revery, M. (1981) *Biochim. Biophys. Acta* **660,** 148–150.
22. Docherty, A. J. P., Bodmer, M. W., Angal, A., Verger, R., Riviere, C., Lowe, P. A., Lyons, A., Emtage, J. S. & Harris, T. J. R. (1985) *Nucleic Acids Res.* **13,** 1891–1903.
23. Weisgraber, K. H., Rall, S. C., Jr., Mahley, R. W., Milne, R. W., Marcel, Y. L. & Sparrow, J. T. (1986) *J. Biol. Chem.* **261,** 2068–2076.
24. Kuusi, T., Bry, K., Nikkilä, E. A. & Kinnunen, P. K. J. (1979) *Medical Biology* **57,** 192–195.
25. Harper, E. (1980) *Annu. Rev. Biochem.* **49,** 1063–1078.
26. Reddy, M. N., Maraganore, J. M., Meredith, S. C., Heinrikson, R. L. & Kezdy, F. J. (1986) *J. Biol. Chem.* **261,** 9678–9683.