Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 2335–2339, April 1986
Biochemistry

Cloning and expression of human lecithin–cholesterol acyltransferase cDNA

(lipid metabolism/cholesterol esterification)

JOHN MCLEAN*, CHRISTOPHER FIELDING†, DENNIS DRAYNA*, HANS DIEPLINGER†, BRADFORD BAER*, WILLIAM KOHR*, WILLIAM HENZEL*, AND RICHARD LAWN*

*Departments of Molecular Biology and Protein Biochemistry, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080; and †Cardiovascular Research Institute and Department of Physiology, University of California, San Francisco, CA 94143

Communicated by Richard J. Havel, December 9, 1985

ABSTRACT **cDNA and genomic cloning has been used to determine the mRNA and amino acid sequence of human plasma lecithin–cholesterol acyltransferase (LCATase; EC 2.3.1.43). The mature protein was found to contain 416 amino acid residues with a hydrophobic leader sequence of 24 amino acids. An unusual feature of the message is that the poly(A) signal AATAAA overlaps the COOH-terminal glutamic acid and stop codons, and the 3′ untranslated region is only 23 bases. The protein itself is distinguished by a number of extended sequences of hydrophobic amino acids, one of which contains a hexapeptide identical with the interfacial binding segment of the active site of pancreatic lipase and is similar to the same site of lingual lipase. The cloned cDNA allows the expression of active LCATase by transfected tissue culture cells.**

Cholesterol of peripheral origin is transported through the plasma for catabolism in the liver. The sequence of reactions involved in this pathway (reverse cholesterol transport) is believed to be important in peripheral cholesterol homeostasis (1). A key component in the process of transport and metabolism of cholesterol in plasma is its esterification by lecithin–cholesterol acyltransferase (LCATase; EC 2.3.1.43). This enzyme, a glycoprotein of apparent M_r ≈63,000 (2, 3), is secreted from the liver into the plasma compartment (4, 5). Under physiological conditions, LCATase catalyzes the transfer of acyl groups from the *sn*2 position of lecithin to the 3-OH of free cholesterol. Apolipoprotein A-I, the major protein of plasma high density lipoprotein, is a potent activator of LCATase activity (6). The conversion of diffusible cholesterol to its insoluble ester form in plasma is important in maintaining a concentration gradient between cell membranes and plasma. When LCATase activity is inhibited *in vitro* or *in vivo*, or is genetically defective, cholesterol is no longer transported to plasma and accumulates in the tissues (7–9). Although the LCATase protein has been purified to homogeneity by several laboratories, little is known of its structure or mechanism of action. Here we report the isolation and sequence analysis of human LCATase cDNA clones, the translated amino acid sequence of the mature LCATase protein and its leader prepeptide, and the expression of enzymatically active LCATase from recombinant DNA plasmids.

METHODS

Protein Purification and Sequence. LCATase was purified from normal human plasma (500 ml) by a modification of procedures reported earlier (10). The ρ 1.21–1.25 g/ml fraction was isolated following preparative centrifugation of plasma in NaBr solution (1.24 g/ml). This was passed through a column (2.5 × 20 cm) of phenyl-agarose (Pharmacia, Uppsala) equilibrated with 3 M NaCl/1 mM EDTA, pH 7.4. After washing with 500 ml of 3 M NaCl/10 mM Tris·HCl, pH 7.4, the phenyl-agarose support was further washed with 0.15 M NaCl/Tris buffer until the OD_{280} of the eluate was <0.05. Remaining adsorbed protein was eluted with distilled water and passed onto a column of DEAE-cellulose (DE-52, Whatman) equilibrated with 10 mM Tris·HCl at pH 7.4. The column was eluted with a gradient of NaCl (0–0.3 M NaCl in 10 mM Tris·HCl, pH 7.4). The fraction containing LCATase was added to a column (1.5 × 4 cm) of hydroxylapatite (Bio-Rad, hydroxylapatite HT) equilibrated with distilled water and eluted with a gradient of 0–5 mM sodium phosphate (pH 6.8) in 0.15 M NaCl. The LCATase-containing fraction was collected. Any remaining apolipoprotein D was removed, if necessary, by passing these column fractions through an immunoaffinity column containing specific polyclonal antibody to apolipoprotein D covalently linked to agarose by the CNBr procedure. The final product was judged to be pure by gradient gel electrophoresis as determined by silver staining.

NH_2-terminal sequence analysis was performed on native LCATase on a prototype gas/liquid-phase sequencer with trimethylamine, phenylisothiocyanate, and trifluoroacetic acid as reagents. Extracted anilinothiazolinone amino acid derivatives were automatically converted to phenylthiohydantoin amino acids with 25% aqueous trifluoroacetic acid and were separated on a Beckman ultrasphere octyl column. Tryptic peptides were sequenced by the gas/liquid-phase sequencer or a modified Beckman model 890B sequencer. Tryptic digestion of LCATase was performed in 0.1 M Tris (pH 8.0) with 0.01% Tween 20 at 37°C for 18 hr with an enzyme-to-substrate ratio of 1:20. The digest was chromatographed on a Synchrom RP-4 column (4.6 mm × 10 cm). The elution solvents were 0.1% trifluoroacetic acid in water (solvent 1) and 1-propanol containing 0.07% trifluoroacetic acid (solvent 2). Peptides were eluted with a linear gradient of from 1% solvent 2 and 99% solvent 1 to 50% solvent 2 and 50% solvent 1 at 25°C with a flow rate of 1.0 ml/min using a Spectra Physics model SP8000 HPLC.

cDNA and Genomic Cloning. Human adult liver cDNA libraries in λgt10 were prepared and provided by Axel Ullrich and Lisa Coussens of Genentech, Inc. (11, 12). Phage were screened on filters with single oligonucleotide probes ranging from 33 to 60 bases, as described in *Results*. DNA from positive phage was isolated by standard procedures and fragments were subcloned into M13 phage vectors for DNA sequencing by dideoxy chain termination (13). All sequences reported in this paper resulted from independent analysis of

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: LCATase, lecithin–cholesterol acyltransferase; bp, base pair(s); SV40, simian virus 40.

both DNA strands. A human genomic library in λCharon 30 (14) was screened with restriction fragments of cDNA clones by standard procedures.

RNA Analysis. RNA transfer blot analysis followed electrophoresis of human liver poly(A)$^+$ RNA in 1.5% agarose gels containing 6% formaldehyde. After transfer, the nitrocellulose filter was hybridized with ^{32}P-labeled LCATase cDNA clone pL12 in 0.75 M NaCl, 75 mM trisodium citrate, 50 mM sodium phosphate (pH 6.8), 5× concentrated Denhardt's solution (15), 50% formamide, 10% dextran sulfate, and 20 μg of boiled, sonicated salmon sperm DNA per ml at 42°C, washed in 15 mM NaCl/1.5 mM sodium citrate/0.1% $NaDodSO_4$ at 65°C, and exposed to x-ray film.

Expression and Assay. An LCATase cDNA expression plasmid was constructed by joining the 5′ portion of clone pL12 and the 3′ portion of clone pL4 at the internal *Pst* I site. The resulting 1.5-kilobase *Eco*RI fragment was ligated into plasmid pgDtruncDHFR (16) after removing the *Eco*RI fragment containing the herpes simplex virus gD coding sequences. COS-7 (monkey kidney) cells at a density of 1.5×10^6 cells per 60-mm dish were rinsed in serum-free Eagle's minimal medium, transfected with plasmid pSVLCAT-1 (4 μg/ml) and DEAE (200 μg/ml) (17) in the same medium for 5 hr at 37°C and 7% CO_2, rinsed in serum-free growth medium, and grown in 2.5 ml of serum-free growth medium for 60 hr. (Serum-free growth medium is medium F-12 supplemented with 5 mg of insulin per ml and 10 mg of transferrin per ml.) Supernatants were removed and assayed immediately. Control cultures of COS-7 cells were subjected to the same transfection and growth protocol, except that pSVLCAT.1 DNA was not included. Culture supernatants were assayed for LCATase activity as described (10). Activity is expressed as picomoles of cholesteryl ester synthesized per milliliter of culture medium per hour.

RESULTS

Partial Protein Sequence, Oligonucleotide Probes, and cDNA Cloning. Amino acid sequence was determined for the first 18 NH_2-terminal residues of LCATase and for tryptic digest peptides ranging from 11 to 20 residues. Utilizing tables of codon usage of mammalian genes and seeking to avoid purine–purine mismatches if no clear choice was evident, we designed single oligonucleotide probes of 33–60 bases (see Fig. 1). About 2 million phage from an oligo(dT)-primed human adult liver cDNA library in λgt10 were grown on 40 15-cm Petri plates from which triplicate nitrocellulose filters were lifted. The filters were hybridized with different ^{32}P-end-labeled oligonucleotide probes in 0.75 M NaCl, 75 mM trisodium citrate, 50 mM sodium phosphate (pH 6.8), 5× concentrated Denhardt's solution, 20% formamide, 10% dextran sulfate, and 20 μg of boiled, sonicated salmon sperm DNA per ml at 42°C overnight and washed for 2 hr in 0.15 M NaCl/15 mM trisodium citrate/0.1% $NaDodSO_4$ at 43°C. Nineteen very strongly hybridizing duplicate positives were observed with filters hybridized with the LCAT.4 probe and with filters hybridized with both the LCAT.3 and LCAT.5 probes. Southern blots in these same hybridization solutions revealed ≈10 discrete bands and RNA transfer blots revealed a hybridizing RNA smaller than 18 S. Upon rescreening with separated probes, 15 of 19 picked plaque regions rehybridized with LCAT.4, 12 also with LCAT.3, and 4 with LCAT.1, the probe based on NH_2-terminal protein sequence. The four phage that hybridized with the NH_2-terminal probe were the best candidates for full-length cDNA clones and were subsequently plaque-purified and analyzed by DNA sequencing (along with a fifth cDNA clone recovered subsequently).

cDNA Sequence. The extent of the five LCATase cDNA clones that were sequenced and the deduced structure of

NH_2 terminus of LCATase: LCAT.1

Peptide sequence			Phe	Trp	Leu	Leu	Asn	Val	Leu	Phe	Pro	Pro	His	Thr	Thr	Pro	Lys
Oligonucleotide probe	TTC	TGG	CTG	CTG	AAC	GTG	CTG	TTC	CCT	CCT	CAC	ACC	ACC	CCT	AAG		
Corresponding LCATase sequence	TTC	TGG	CTC	CTC	AAT	GTG	CTC	TTC	CCC	CCG	CAC	ACC	ACG	CCC	AAG		

Tryptic peptide: LCAT.3

Leu	Glu	Pro	Gly	Gln	Gln	Glu	Glu	Tyr	Tyr	Arg	
CTG	GAA	CCT	GGC	CAG	CAG	GAG	GAG	TAC	TAC	CGG	
CTG	GAG	CCC	GGC	CAG	CAG	GAG	GAG	TAC	TAC	CGC	

Tryptic peptide: LCAT.4

Ile	Thr	Thr	Thr	Ser	Pro	Trp	Met	Phe	Pro	Asp	Arg
ATC	ACC	ACC	ACC	TCC	CCT	TGG	ATG	TTC	CCT	GAC	CGG
ATA	ACC	ACC	ACC	TCC	CCC	TGG	ATG	TTT	CCC	TCT	CGC
										(Ser)	

Tryptic peptide: LCAT.5

Gln	Pro	Gln	Pro	Val	His	Leu	Leu	Pro	Ala	His	Gly	Ile	Gln	His	Leu	Asn	Met	Val	Phe
CAG	CCT	CAG	CCT	GTG	CAC	CTG	CTG	CCT	GCT	CAC	GGC	ATC	CAG	CAC	CTG	AAC	ATG	GTG	TTC
CAG	CCA	CAG	CCT	GTG	CAC	CTG	CTG	CCC	CTG	CAC	GGG	ATA	CAG	CAT	CTC	AAC	ATG	GTC	TTC
									(Leu)										

FIG. 1. Peptide sequence and oligonucleotides probes. In each of four cases, the direct peptide sequence analysis (top line) led to the synthesis of the unique "long" oligonucleotide probes shown. (The actual probes were the reverse complement of these sequences so that they could be used for hybridization to RNA as well as to DNA.) The corresponding LCATase cDNA sequences are shown on the bottom line. In two cases, an amino acid coded by the cDNA differed from that predicted from peptide sequencing (as indicated by parentheses).

LCATase mRNA are shown in Fig. 2. Fig. 3 displays the DNA sequence and translated sequence of the LCATase cDNA. The DNA sequences of all five LCATase cDNA clones were similar, differing only in the extent of LCATase mRNA represented and in the length of poly(A) tails. The two longest clones, pL4 and pL12, were sequenced to completion, whereas cytosine reactions only were performed for pL2, pL10, and pL19. Clones pL2, pL4, pL10, and pL19 contained poly(A) tails of different lengths (50–90 bases) at the same location, whereas pL12 did not extend back to the poly(A) tail. Clones pL4 and pL12 extended farthest at the 5′ end. They both contained the entire coding sequence for the LCATase protein but contained only short regions of 5′ untranslated sequence. The DNA sequence of the independent clones pL4 and pL12 were identical throughout the ≈1400 base pairs (bp) for which they correspond, serving to confirm the veracity of the cloning and sequencing procedures.

The 5′ end of the cDNA contains an initiator methionine codon followed by a continuous open reading frame that codes for a 440 amino acid polypeptide. The first 24 residues contain a core of hydrophobic amino acids and are likely to represent an NH_2-terminal secretion signal peptide. Thereafter follows the sequence (beginning Phe-Trp-Leu) of the NH_2 terminus of the LCATase protein as it is purified from plasma. The mature protein contains 416 amino acids with a calculated M_r of 47,090. LCATase is known to be a glycoprotein that migrates on $NaDodSO_4$/polyacrylamide gel at M_r ≈63,000. Previous authors estimated a carbohydrate content of 25% and a deduced polypeptide weight of about M_r 45,000 (2). The translated DNA sequence predicts four possible N-linked glycosylation sites (Asn-Xaa-Ser; Asn-Xaa-Thr). Glycosylation at one of these sites (residue 272) was detected during the course of peptide sequencing. Other potential sites of glycosylation remain unconfirmed.

The methionine codon at nucleotides 12–14 (Fig. 3*a*) is presumed to initiate transcription of the preprotein. This ATG is followed by guanosine and preceded by guanosine at −3 nucleotides, in reasonable agreement with consensus sequences near the translation start of eukaryotic mRNA (19). None of the cDNA clones we characterized contains the entire 5′ untranslated region of the message. RNA transfer blot hybridization of human liver poly(A)$^+$ RNA probed with LCATase cDNA clone 12 (Fig. 2*a*) reveals a single hybrid-

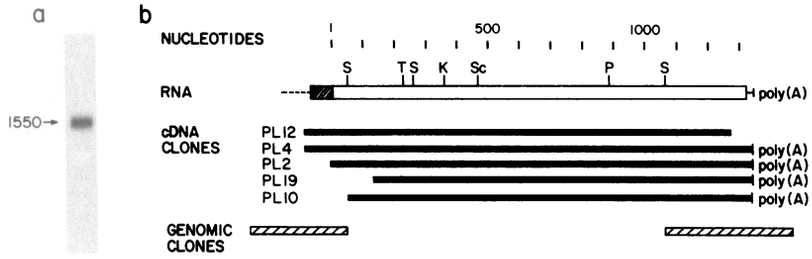

FIG. 2. Human LCATase mRNA. (*a*) RNA transfer blot hybridization. Five micrograms of poly(A)$^+$ RNA from human liver was electrophoresed in a 1.5% agarose/6% formaldehyde-containing gel, transferred to nitrocellulose, and probed with ^{32}P-labeled LCATase cDNA clone pL12. The single hybridizing band measures 1550 ± 50 nucleotides. *E. coli*, yeast, and bovine rRNAs and *Hae* III-digested ϕX DNA served as size standards. (*b*) Human LCATase mRNA, as deduced from cDNA clones, is depicted below a size scale in nucleotides. Direction of transcription is from left to right. Features of the message are 5′ untranslated region (estimated size; dotted line), leader peptide coding region (hatched box), mature coding region (open box), and 3′ untranslated region ending in poly(A) site (line with bar). All of the sites for the restriction enzymes S (*Sst* I), T (*Taq* I), K (*Kpn* I), Sc (*Sca* I), and P (*Pst* I) are shown. Below this line, the five horizontal bars indicate the extent of characterized cDNA clones. Both strands of clones pL4 and pL12 were sequenced to completion, whereas the extents of clones pL2, pL10, and pL19 were subject to cytosine reactions only. Hatched bars on the bottom line depict regions of genomic clones that were sequenced to elucidate the terminal regions of the gene.

izing band of 1550 ± 50 bases. This implies that LCATase mRNA contains ≈100 bases of 5′ untranslated sequence.

The 3′ end of the cDNA clone sequences reveals an unusually short 3′ untranslated region of 23 nucleotides. Indeed, the common poly(A) signal AATAAA that precedes eukaryotic poly(A) sites by 20–30 nucleotides (20) is partially contained in the codon for the COOH-terminal glutamine (GAA) and in the translation stop codon TAA. The four independent poly(A)-containing clones we analyzed all had poly(A) tails at the same site. RNA transfer blot analysis did not indicate the abundant presence of longer RNA species, suggesting that the proposed location is the major site of polyadenylylation.

Genomic Clones. Due to the incomplete 5′ ends and the unusual features of the 3′ ends of LCATase cDNA clones, we isolated and analyzed the corresponding regions of genomic DNA. A *Sau*3AI partial-digest library of human genomic DNA in phage λCharon 30 (14) was screened with 3′ and 5′ terminal *Sst* I fragments from pL12. Five hybridizing clones were recovered and all appeared to contain the entire cDNA coding sequence, as determined from restriction mapping and hybridization experiments. Phage DNA from one of these genomic clones, λL1, was cut with several restriction enzymes recognizing five or fewer bases, transferred to nitrocellulose by the method of Southern, and hybridized with the terminal probes. Both probes hybridized to *Alu* I fragments ≈400 bp in length. *Alu* I-digested λL1 DNA in this size range was isolated by gel electrophoresis and cloned into bacteriophage M13. Plaques hybridizing to terminal *Sst* I probes were isolated and subjected to dideoxy sequencing, resulting in the sequence of genomic fragments overlapping the 5′ and 3′ ends of LCATase cDNA (Fig. 3*b*).

The genomic sequence extends 267 bp 5′ of the methionine codon presumed to initiate translation of the preprotein. No ATG triplets are found in this region, whereas stop codons occur in all three reading frames. The 3′ genomic sequence extends 154 bp beyond the stop codon and agrees with the sequence of the short 3′ untranslated region and the poly(A) site inferred from the cDNA clones. No other AATAAA sequences, nor recognized variants of this poly(A) signal, occur in this region of the genome, which extends 134 bp beyond the poly(A) site in the cDNA clones.

Expression of the Cloned LCATase Gene. The full-length LCATase cDNA was assembled from the clones pL4 and pL12 and inserted into an expression plasmid containing the SV40 origin of replication and early promoter to effect expression of LCATase coding sequences. The plasmid is designated pSVLCAT.1. COS-7 monkey kidney cells were transfected with this DNA and the culture supernatants were assayed for LCATase activity after 60 hr. Background activities in serum-free growth medium and in medium from mock-transfected COS-7 cells were also determined. Based on three separate experiments, LCATase activity in the medium of transfected cells was 4 ± 1.9 pmol·ml^{-1}·hr^{-1}, whereas activity in the medium of control cells cultured under the same conditions was 0.5 ± 0.4 pmol·ml^{-1}·hr^{-1}. There was no detectable activity in the medium when 1.5 mM 5,5′-dithiobis(2-nitrobenzoic acid), a known LCATase inhibitor (21), was included in the assay medium or when apolipoprotein A-I was not present. These data indicate that the appearance of LCATase in the culture medium is induced after transfection with plasmid pSVLCAT-1 and that this activity has the properties of the plasma enzyme in terms of its cofactor dependence and inhibition by sulfhydryl reagents. These results are derived from a transient expression system and serve to verify that the cloned cDNA codes for LCATase, which can be expressed in heterologous cells as a recombinant DNA-derived product. Eventual large-scale production of LCATase will clearly depend on expression in microorganisms or tissue culture systems from stably incorporated recombinant DNA.

DISCUSSION

Full-length or near full-length LCATase cDNA clones were efficiently recovered from a λgt10 library utilizing unique sequence oligonucleotide probes representing one codon choice for each of four different stretches of amino acid sequence. As demonstrated previously (14, 22, 23), the use of unique sequence "long probes" has several advantages over the use of pools of 14- to 17-nucleotide probes synthesized to represent all possible codon choices. The number of false positive clones can be reduced, and single long probes are less sensitive to errors in protein sequencing. Two of the probes used here were successful, even though one amino acid prediction in each failed to match the cloned cDNA sequences. The use of several independent probes increases the proportion of desired clones among hybridization positives, and "previewing" the probes with genomic DNA Southern blots and RNA transfer blots aids in the evaluation of the probe and in determining optimal hybridization conditions. In the current case, the inclusion of a probe based on NH_2-terminal sequence allowed the rapid identification of full-length clones in the initial screen. This approach has the added advantage of streamlining the efforts in protein sequencing. An efficient strategy is to obtain NH_2-terminal sequence on intact protein as well as sequence from several

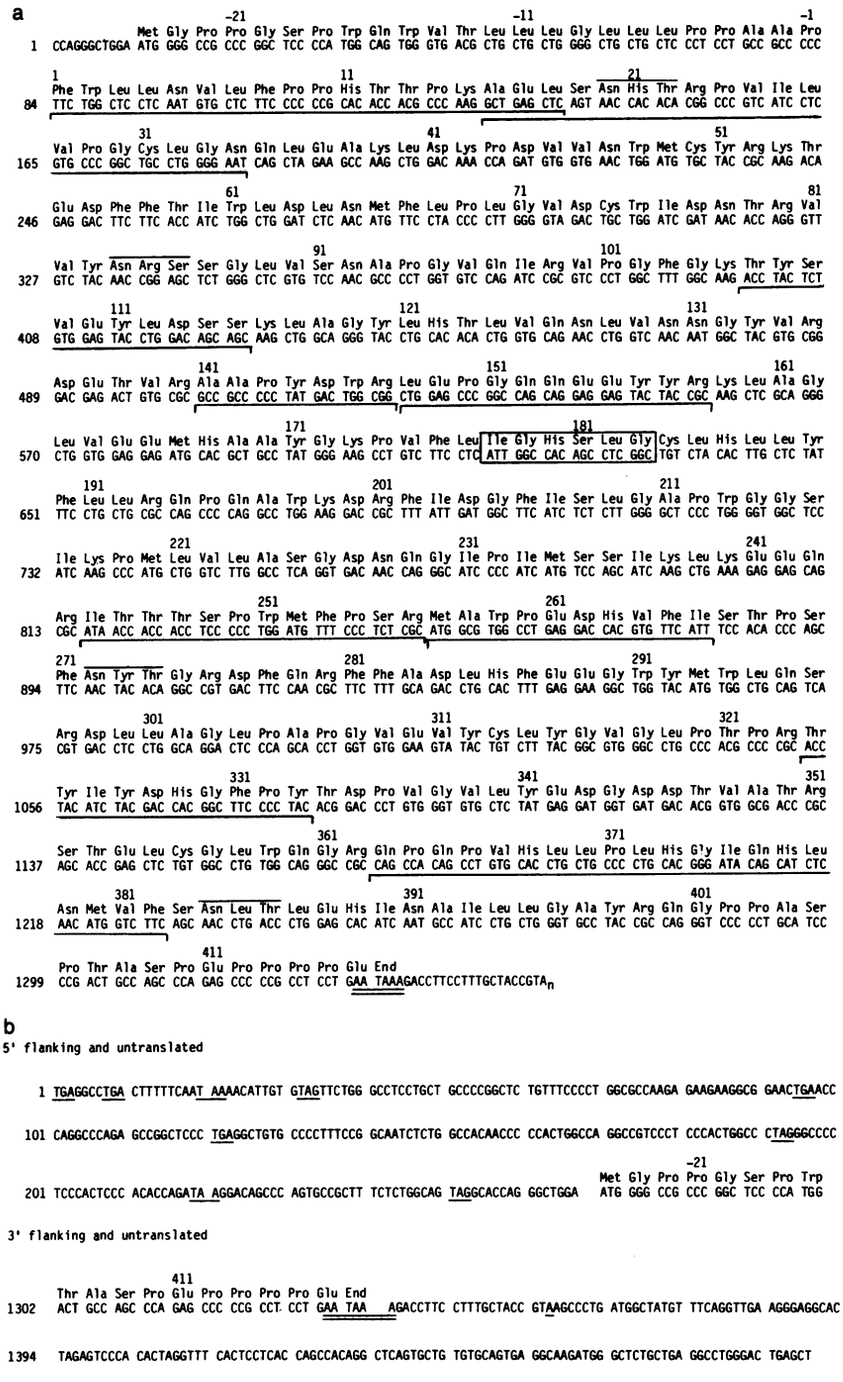

FIG. 3. DNA sequence analysis. (*a*) Human LCATase cDNA. Nucleotides (left of each row) are numbered from the 5′ terminus of cDNA clone pL12. The complete predicted amino acid sequence of LCATase is shown above the DNA sequence. Negative amino acid numbers (above residues) refer to the presumed leader prepeptide; positive numbers refer to the mature protein. Eighteen NH_2-terminal residues sequenced here, in agreement with eight reported by Doi and Nishida (18), serve to position the mature NH_2 terminus. The four predicted N-linked glycosylation sites are overlined, and the six-residue stretch (surrounded by hydrophilic regions), which is identical to the interfacial binding site of porcine pancreatic lipase, is boxed in. Peptide sequence determined directly from the NH_2 terminus or from tryptic fragments is underlined; some residues in these regions were ambiguous. Double underlines indicate the conserved poly(A) signal hexanucleotide. Poly(A) tails of varying length occur at the same location in four cDNA clones. Complete DNA sequencing was performed on both strands of two independent clones by dideoxy chain-termination procedures. (*b*) Genomic sequences obtained from bacteriophage λ clones are shown that overlap the 5′ and 3′ ends of the cDNA, respectively. 5′ Flanking nucleotides are arbitrarily numbered from the 5′ end of available sequence. NH_2-terminal LCATase protein sequence is numbered as in *a*. Stop codons in all three reading frames are underlined. The first in-frame upstream stop codon is the TGA of nucleotides 121–124. 3′ Flanking nucleotides and COOH-terminal peptides are numbered as in *a*. The poly(A) signal hexanucleotide has double underlines and the site of poly(A) tail addition has single underlines.

tryptic peptides and to use all of these as the basis of unique sequence long probes. There is no demand for the often laborious search for sequences of low codon redundancy.

The human LCATase mRNA is ≈1550 nucleotides, as revealed by RNA transfer blot hybridization. Together with the analysis of cDNA and genomic clones, this implies that the 1320-base open reading frame is flanked by ≈100 bases of 5′ untranslated sequence. The longest cDNA clone extends only 11 bp of 5′ of the ATG believed to initiate translation. However, the sequence from genomic clones (Fig. 3*b*) reveals no other ATG triplets in the 267 bp 5′ to the ATG start codon, whereas stop codons occur in all possible reading frames within 147 bases of this ATG. We have not yet mapped the true 5′ terminus of the LCATase mRNA nor determined whether introns are located in the 5′ untranslated region.

The 3′ untranslated region of the LCATase gene is unusually short. Only 23 nucleotides separate the final glutamine codon from the poly(A) tail. Indeed, the ubiquitous poly(A) signal AATAAA is partially contained within the glutamine and stop codons. To our knowledge, such parsimony of RNA sequence has been reported in only two other eukaryotic nuclear genes (31, 32). The site identified appears to be the principal site of polyadenylylation of the LCATase message in human liver for several reasons. (*i*) All of the four independent cDNA clones that we sequenced contained poly(A) tails at the same location. (*ii*) DNA sequence of the genomic LCAT clone revealed no other AATAAA (or closely related) poly(A) signal sequences within the 134 further 3′ nucleotides analyzed (Fig. 3*b*). (*iii*) The genomic sequence contains the octanucleotide TATGTTTC beyond the poly(A) site that conforms to a consensus sequence located downstream from 67% of mammalian mRNA termini examined

(24). (*iv*) Blot hybridization (Fig. 2*a*) revealed a single hybridizing RNA species of ≈1550 bases without other noticeable discrete larger forms. Poly(A) tails of ≈100 nucleotides at this location and a 5′ untranslated region of ≈100 bases are consistent with this mRNA size determination and an open reading frame of 1320 bases.

Translation of the cloned cDNA sequence provides the complete amino acid sequence of LCATase. The mature protein of 416 residues is preceded by a hydrophobic leader sequence of 24 amino acids. The calculated M_r of this polypeptide agrees with previous size estimates for a glycoprotein of M_r ≈60,000–67,000, which is believed to contain ≈25% carbohydrate by mass. The amino acid composition predicted from the cDNA sequences falls within the range of values from five independent determinations, except that we predict six cysteines, rather than the four reported by amino acid analysis by Chong *et al.* (25). If the first 24 amino acids of the open reading frame do indeed represent a cleaved signal sequence, then the mature protein begins with the sequence Phe-Trp-Leu-Leu, etc., in agreement with NH_2-terminal sequence of the plasma-derived protein determined here as well as the eight NH_2-terminal residues reported by Doi and Nishida (18). These agreements with previous reports, as well as the expression of LCATase activity from the plasmid pSVLCAT.1 and the sequence homology discussed below, support the conclusion that the cDNA clone described here codes for LCATase.

Although no other cholesterol acyltransferase has been sequenced, the human LCATase sequence shares some features with sequences of other catalytic factors that interact with lipids. Several enzymes and transfer proteins have been found to bind lipids at "interfacial" active sites consisting of extended linear sequences of hydrophobic amino acids (26–28). The amino acid sequence of LCATase conforms to this pattern. In addition to the overall hydrophobic character of the protein, several linear sequences of five or more hydrophobic residues can be identified. Of special interest are the sequences -Val-Phe-Leu-Ile-Gly- and -Leu-Gly-Cys-Leu-His-Leu-Leu-Tyr-Phe-Leu-Leu-, which surround the hexapeptide -Ile-Gly-His-Ser-Leu-Gly- (residues 178–183). This sequence is identical to residues 149–154 of porcine pancreatic lipase. Furthermore, the serine in this hexapeptide has been identified as part of the interfacial active site of that lipase (29). Rat lingual triglyceride lipase has been found to contain a similar active site sequence (-Val-Gly-His-Ser-Gln-Gly-) (30). These data are suggestive of a common hydrophobic attachment site in several different enzymes of lipid metabolism. There is no other substantial homology between them, however, and whether the similarity observed is the result of convergence or a common origin cannot presently be determined.

To our knowledge, aside from LCATase, no other enzymes of plasma lipid metabolism have been sequenced. In their reactions with plasma lipids in lipoprotein form, these enzymes often show complex requirements of cofactor dependence and reactivity strongly dependent upon the composition of the lipoprotein surface. The availability of the sequence of LCATase may represent a first step in determining the molecular basis of the specificity and mode of action of this enzyme.

We thank Parkash Jhurani and Peter Ng for oligonucleotide synthesis, Ellson Chen and Lucy Rhee for assistance in DNA sequence analysis, and Jeanne Arch and Alane Gray for preparation of the manuscript and illustrations, respectively. This work was supported by Genentech, Inc., by Arteriosclerosis SCOR Grant HL14237 to C.F., and by a Max Kade Foundation grant to H.D.

1. Fielding, C. J. & Fielding, P. E. (1980) *Med. Clin. N. Am.* **66,** 363–373.
2. Chung, J., Abano, D. A., Fless, G. M. & Scanu, A. M. (1979) *J. Biol. Chem.* **254,** 7456–7464.
3. Doi, Y. & Nishida, T. (1983) *J. Biol. Chem.* **258,** 5840–5846.
4. Osuga, T. & Portman, O. W. (1970) *Am. J. Physiol.* **220,** 735–741.
5. de Parscau, L. & Fielding, P. E. (1984) *J. Lipid Res.* **25,** 721–728.
6. Fielding, C. J., Shore, V. G. & Fielding, P. E. (1972) *Biochim. Biophys. Res. Commun.* **46,** 1493–1498.
7. Fielding, C. J. & Fielding, P. E. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3911–3914.
8. Davis, R. A., Halgerud, P., Dueland, S. & Drevon, C. A. (1982) *Biochim. Biophys. Acta* **689,** 410–414.
9. Glomset, J. A. & Norum, K. R. (1973) *Adv. Lipid Res.* **11,** 1–65.
10. Aron, L., Jones, S. & Fielding, C. J. (1978) *J. Biol. Chem.* **253,** 7220–7226.
11. Ullrich, A., Coussens, L., Hayflick, J., Dull, T., Gray, A., Tam, A., Lee, J., Yarden, Y., Liberman, T., Schlessinger, J., Downward, J., Myers, E., Whittle, N., Waterfield, M. & Seeburg, P. (1984) *Nature (London)* **309,** 418–425.
12. Huynh, T., Young, R. & Davis, R. (1985) in *DNA Cloning, A Practical Approach*, ed. Glover, D. (IRL, Oxford), Vol. 1, pp. 49–78.
13. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9,** 309–321.
14. Wood, W. I., Capon, D., Simonsen, C., Eaton, D., Gittschier, J., Keyt, B., Seeburg, P., Smith, D., Hollingshead, P., Wion, K., Delwart, E., Tuddenham, E., Vehar, G. & Lawn, R. (1984) *Nature (London)* **312,** 330–337.
15. Denhardt, O. (1966) *Biochem. Biophys. Res. Commun.* **23,** 641–646.
16. Lasky, L. A., Dowbenko, D., Simonsen, C. C. & Berman, P. W. (1984) *Biotechnology* **2,** 527–532.
17. McCutchan, J. H. & Pagano, J. (1968) *J. Natl. Cancer Inst.* **41,** 351–357.
18. Doi, Y. & Nishida, T. (1981) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **40,** 1695 (abstr.).
19. Kozak, M. (1980) *Nucleic Acids Res.* **8,** 127–142.
20. Proudfoot, N. & Brownlee, G. (1981) *Nature (London)* **252,** 359–362.
21. Stokke, K. T. & Norum, K. R. (1971) *Scand. J. Clin. Lab. Invest.* **27,** 21–27.
22. Anderson, S. & Kingston, I. B. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 6836–6842.
23. Ullrich, A., Berman, C. H., Dull, T. J., Gray, A. & Lee, J. M. (1984) *EMBO J.* **3,** 361–364.
24. McLauchlan, J., Gaffney, D., Whitton, J. & Clements, J. (1985) *Nucleic Acids Res.* **13,** 1347–1368.
25. Chong, K. S., Jahani, M., Hara, S. & Lacko, A. G. (1983) *Can. J. Biochem. Cell Biol.* **61,** 875–881.
26. Pujik, W. C., Verhij, H. M. & de Haas, G. H. (1977) *Biochim. Biophys. Acta* **492,** 254–259.
27. de Caro, J., Boudouard, M., Bonicel, J., Guidoni, P., Desnuelle, P. & Rovery, M. (1981) *Biochim. Biophys. Acta* **671,** 129–138.
28. Akeryod, R., Moonen, P., Westerman, J., Puyk, W. & Wirtz, K. (1981) *Eur. J. Biochem.* **114,** 385–391.
29. Guidoni, A., Benkouka, F., deCaro, J. & Rovery, M. (1981) *Biochim. Biophys. Acta* **660,** 148–150.
30. Docherty, A. J. P., Bodmer, M., Angal, S., Verger, R., Riviere, C., Lowe, P., Lyons, A., Emtage, J. & Harris, T. (1985) *Nucleic Acids Res.* **13,** 1891–1903.
31. Fiddes, J. C. & Goodman, H. M. (1980) *Nature (London)* **286,** 684–687.
32. Leytus, S., Chung, D., Kisiel, W., Kurachi, K. & Davie, E. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3699–3702.