Cloning and expression of apolipoprotein B, the major protein of low and very low density lipoproteins

(Agt11 expression system/immunochemical identification/large mRNA)


*Departments of Medicine and Microbiology, and ‡Molecular Biology Institute, University of California, Los Angeles, CA 90024; and †Research Service, Veterans Administration, Wadsworth Medical Center, Los Angeles, CA 90073

Communicated by Daniel Steinberg, March 12, 1985

ABSTRACT We report the cloning of cDNAs for rat liver apolipoprotein B (apo B) and the use of the cloned sequences to examine apo B expression at the level of mRNA in rat tissues. Fifteen putative apo B clones were identified by antibody screening of a rat liver cDNA library in the Agt11 expression vector. The identity of the clones was confirmed by immunological studies of the fusion protein products. All clones appear to contain sequences found only in apo B PI, the high molecular weight form of rat liver apo B. Blotting studies show that the clones hybridize to a single 20-kilobase liver mRNA species, sufficiently large to encode the entire apo B PI peptide, which is estimated to be 400 kDa in size. Apo B PI mRNA is abundant in liver and present in lower amounts in intestine but is absent in a variety of other tissues examined. This tissue distribution is consistent with that expected from studies on the in vivo synthesis of apo B. One clone, corresponding to a 240-base segment of the apo B PI mRNA, was sequenced and found to exhibit homology with a short region of rat apo E mRNA. Analysis of the secondary structure of the corresponding peptide did not show the preponderance of amphipathic α-helical structures characteristic of other apolipoproteins examined thus far.

Apolipoprotein B (apo B) is a major protein component of mammalian very low density lipoproteins (VLDL) and chylomicrons and essentially the sole protein in low density lipoproteins (LDL). It is essential for the assembly/secretion of chylomicrons and VLDL, since these lipoproteins are absent in individuals with abetalipoproteinemia, a genetic disorder in which apo B is not produced (1). Apo B also functions as the ligand for the removal of LDL from the circulation by receptor-mediated uptake into a variety of cells (2, 3).

Despite intensive study, the structure of apo B has remained an enigma. It is among the largest peptides known, with estimates ranging from 250 to 550 kDa (1, 4-10), and is extremely hydrophobic as well as very sensitive to proteolytic degradation, and only recently has limited amino acid sequence information become available (11). Moreover, at least two major forms of apo B are present in the circulation (8, 9, 12-15). In humans and most other mammals, the liver incorporates a 400-kDa species into VLDL, whereas the intestine incorporates a 210-kDa species into chylomicrons. However, the rat and mouse are exceptional in that their livers produce VLDL with comparable amounts of both these peptides (9, 13-15). We call these high and low molecular weight forms of apo B P(epsilon)I and P(epsilon)III, respectively; they are the B-100 and B-48 species in Kane's nomenclature (1, 8, 9). Although in the rat, PI is also accompanied by a slightly smaller component, PII, the two are very similar (9) and will be collectively referred to as PI. Peptide fingerprinting and immunological analysis suggest that PI encompasses a unique, PI-specific moiety joined to a second moiety which is identical or very similar to PII (9, 16-18). The possibility that the two peptides share common exons is also supported by the fact that most abetalipoproteinemic patients lack both PI and PII (19). Since our polynucleotide run-off translation experiments argue against proteolytic processing of PI to PIII (unpublished data), this would most simply be accounted for by having separate PI and PII mRNAs formed through differential processing of a single primary apo B transcript.

We have sought to clone nucleic acid sequences encoding the apo B peptides, in order to determine their amino acid sequences and to examine the origin of the various molecular weight forms of apo B. These efforts have been hampered by our inability to translate apo B mRNA either in vitro or in frog oocytes (unpublished) and by the scant amino acid sequence information available. Thus, attempts to obtain apo B clones from liver polysomes enriched for apo B sequences by anti-apo B immunelection were unsuccessful, as was screening of cDNA libraries with mixed oligonucleotide probes. We now report the isolation of rat PI-specific cDNA clones identified by antibody screening of a rat liver cDNA library in the phage Agt11 expression system (20, 21). Results of RNA blot-hybridization analysis with these probes are consistent with the translation of the apo B peptides from separate messages.

MATERIALS AND EXPERIMENTAL PROCEDURES

Materials. Bacterial strains and a library of rat liver poly(A)⁺ RNA cDNA in Agt11 (22) were kindly provided by J. Schwarzbauer (Massachusetts Institute of Technology). The average insert size was estimated to be 1 kilobase pair (kbp). Radioisotopes and restriction enzymes were from Amersham and New England Biolabs, respectively.

Antibodies. Anti-PI. Scatchard plot analysis of rabbit antiserum raised against highly purified rat apo B PI (unpublished data) showed the presence of two major components: "PI-specific" antibodies, with high affinity (Kd = 10⁻¹⁰ M) for epitopes found on PI but not on PII, and "PI/PII-common" antibodies with an order-of-magnitude-lower average affinity for epitopes present on both PI and PII.

Abbreviations: LDL, low density lipoproteins; VLDL, very low density lipoproteins; apo, apolipoprotein; IPTG, isopropyl β-D-thiogalactoside; pb, base pair(s); kb, kilobase(s).

† To whom reprint requests should be addressed.

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.
Affinity-purified anti-PI. Anti-PI was passed over a column of immobilized PI, which had been homogenized by preparative NaDodSO4/PAGE and high performance gel exclusion chromatography. The flow-through was used as PI-adSORbed control serum; bound antibodies were eluted with 0.1 M glycine (pH 3) and designated affinity-purified anti-PI.

PI-specific antibodies. PI/PIII-common antibodies were removed from the anti-PI antisem by passing it over a column of immobilized pure PI. The flow-through was devoid of reactivity with standard PIII but retained its reactivity with standard PI.

Antibody Screening of Rat Liver Ygt11 Library. About 250,000 plaque-forming units (pfu) of recombinant phage were screened with anti-PI antisem diluted 1:100 in Tris/NaCl/3% BSA [10 mM Tris Cl, pH 7.5/0.15 M NaCl containing 3% bovine serum albumin], at a density of 5000 pfu per 85-mm plate of Escherichia coli strain Y1090 (20, 21).

Immunological Characterization of Fusion Proteins. Lysogeny was confirmed for infected E. coli Y1089 cells by colony growth at 32°C but not at 42°C. Lysogens were grown from single colonies in LB medium to OD600 = 0.5 at 32°C, shifted to 45°C for 60 min to induce phage growth, and incubated for another 60 min at 37°C in the presence of the β-galactosidase inducer isopropyl β-D-thiogalactoside (IPTG). Cells were harvested by centrifugation at room temperature, suspended in 10 volumes of Tris/NaCl, disrupted by sonication on ice, and centrifuged to remove insoluble material. Proteins were subjected to NaDodSO4/PAGE in 4%-12% gradient minigels and then electrophoretically transferred to nitrocellulose. Filters were blocked with Tris/NaCl/3% BSA, incubated with the appropriate antibody overnight at room temperature, washed with Tris/NaCl, and developed with 125I-labeled protein A/autoradiography or with an ELISA procedure.

RNA Isolation, Electrophoresis, and Blot-Hybridization. RNA was prepared from tissues by homogenization in guanidinium thiocyanate followed by centrifugation over a cesium chloride cushion (23); poly(A) RNA was enriched by oligo(dT)-cellulose chromatography. RNA was electrophoresed in 1.0% agarose gels in the presence of formaldehyde, transferred to nitrocellulose, and hybridized to 32P-labeled isolated cDNA inserts (specific activity about 108 cpm/μg) for 1 day at 42°C in 50% (vol/vol) formamide/10% (wt/vol) dextran sulfate/0.15 M NaCl/15 mM sodium citrate, pH 7.4/1 mM EDTA/0.1% NaDodSO4/salmon sperm DNA (100 μg/ml). Blots were washed in 15 mM sodium citrate (pH 7.4) at 68°C and autoradiographed. Hybrid-selection and translation were as described (24).

Subcloning and DNA Sequencing. cDNA inserts, isolated by EcoRI restriction of λ phage DNA followed by electrophoresis through agarase and electrophoretic elution, were subcloned in pBR322 and M13. Plasmids and single-stranded phage DNA were isolated as described (25). Sequencing was carried out by the dideoxy chain-termination method of Sanger et al. (26), using synthetic oligonucleotide primers. Computer analysis of DNA sequences used the programs and sequence files of Staden (27), Lipman and Wilbur (28), and Orcut and Dayhoff (29).

RESULTS

Isolation of Apo B cDNA Clones. Initial plaque-screening of the rat liver cDNA library in Ygt11 was performed with 250,000 independent recombinants, using rabbit antibodies against pure rat apo B PI. This antiserum contains high-affinity antibodies that recognize epitopes specific for PI and lower-affinity antibodies that recognize epitopes common to both PI and PIII; thus, under the conditions used for the plaque screens, the detection of PI (about 100 pg) was an order of magnitude more sensitive than that for PIII. A total of 15 independent putative clones were isolated after three or more rounds of purification, and their identity was confirmed by immunological and RNA hybridization studies.

Immunological Characterization of Fusion Proteins. The interaction of induced fusion protein products with antibodies against rat apo B PI or E. coli β-galactosidase on immunoblots is shown in Fig. 1. Upon probing with anti-β-galactosidase the nonrecombinant Ygt11 lysogen gave the expected major 116-kDa β-galactosidase band, whereas the recombinant apo B clones 2 and 49a consistently showed several bands (Fig. 1A). Of these, only the largest species, with apparent molecular mass 130 and 122 kDa, respectively, also reacted with the antibodies against rat apo B PI as

![Fig. 1](image-url). Immunoblotting of fusion proteins induced in two apo B clones. Shown is a composite of blots probed with anti-β-galactosidase (1:2000; Cappelli) (A), anti-PI (1:1000) (B), PI-adSORbed anti-PI (1:1000) (C), affinity-purified anti-PI (1:20) (D), and PI-specific antibody (1:1000) (E). Samples analyzed were 5 μg of PI plus 9 μg of PIIII (lane 1'), 100 ng of PI plus 100 ng of PIIII (lanes 1), 50 μg (based on protein analysis) of IPTG-induced clone 2 lysate (lanes 2), 50 μg of IPTG-induced clone 49a lysate (lanes 3), 50 μg of IPTG-induced nonrecombinant Ygt11 lysogen lysates (lanes λ), and 100 ng of β-galactosidase (lane β). Reactive bands were visualized with an ELISA system (Cappelli) (A) or by 125I-labeled protein A/autoradiography (B-E).
expected for the intact fusion proteins (Fig. 1B). Thus, the abundant species at about 95 kDa and the variable component very close to β-galactosidase presumably represent degraded fusion protein fragments that have lost their apo B portions through proteolytic cleavage. Preliminary experiments indicate that this degradation occurs in vivo, a common problem with fusion proteins produced in E. coli, although the insertion site in λgt11 and the use of the lon host were designed to minimize this problem (20, 21). Synthesis of both β-galactosidase and fusion proteins was dependent on induction with IPTG (data not shown). The reactivity of the fusion proteins or the apo B standards with anti-PI was unaffected by preabsorption of the antisera with rat plasma freed of lipoproteins by ultracentrifugation at d > 1.21 g/ml (data not shown). In contrast, the same amounts of rat apo B standards and intact fusion proteins failed to react with the anti-PI antiserum after it had been adsorbed with immobilized, highly purified PI (Fig. 1C), while retaining reactivity with the affinity-purified anti-PI antibody subsequently eluted from the immobilized antigen (Fig. 1D).

Since the anti-PI antibody used in the original screening reacted preferentially with PI as compared to PIII, it was expected that the clones might encode PI-specific antigens. This was indeed confirmed: antiserum that had been rendered specific for PI by adsorption with PIII maintained its reactivity with the fusion protein from clone 2 (Fig. 1E); the same was also true for clone 49a, as well as for the plaque assays of the complete panel of positive clones (data not shown).

Compared to the 116-kDa β-galactosidase peptide, the apparent sizes of the intact fusion proteins from clones 2 and 49a (130 and 122 kDa, respectively) are those expected for the maximum additional 14- and 6-kDa apo B peptide inserts that could be encoded by their 560- and 240-bp cDNA inserts, respectively.

Characterization of the Cloned cDNA Sequences. Crosshybridization of the isolated inserts showed that, of the 15 apo B clones, 12 contained the identical 240-bp fragment represented by clone 49a, with the remaining clones 2, 29, and 40a (560, 360, and 800 bp, respectively) representing three additional nonoverlapping apo B PI sequences (data not shown).

The sequence of clone 49a (Fig. 2) shows a single open reading frame, continuous with that of lacZ. Chou-Fasman analysis (29) of the corresponding amino acid sequence predicts about 60% α-helix and 15% β-sheet; the α-helices do not have pronounced amphipathic features. A search of the Dayhoff data file showed no extensive homology for the amino acid sequence of clone 49a; however, a short stretch of its nucleotide sequence (bases 62-77 in Fig. 2) is highly homologous (15/16 bases) to a 5′ region of rat apo E mRNA. Dot-matrix analysis (27) shows weak internal repetitive structure (about 40% homology at intervals of about 40 bases). Clone 49a lacked a poly(A) tail, suggesting that synthesis of cDNA was initiated by self-priming or that the 3′ end had been lost during the cloning procedures.

Identification of mRNA for Apo B PI. On blots of rat liver total and poly(A)+ RNA, the nonoverlapping clones 2 and 49a were hybridized to RNA species corresponding to about 20 kilobases (kb) (Fig. 3), with diffuse hybridization in lower molecular weight regions probably representing partial degradation of this uncommonly large species. Since the 400-kDa human PI corresponds to an 11-kb coding sequence, this 20-kb mRNA would be almost twice the size necessary to encode a rat PI peptide. The probes also crosshybridized with a species of similar size in mouse, but not human, liver mRNA (data not shown). In contrast, under the same conditions no hybridization was observed with any component of RNA from rat kidney, heart, brain, or spleen; however, after prolonged autoradiographic exposure, a faint 20-kb band could be seen in the intestinal sample, whereas the other extrahepatic samples remained completely blank. Thus, in the strong hybridization signal for the PI message in rat liver (about one-tenth of that for apo E mRNA in mouse liver (data not shown)), its presence in lower amounts in intestine and its absence in the other extrahepatic tissues is consistent with the known restriction of apo B expression to liver and intestine.

Southern Blotting. Southern blotting analysis of mouse genomic DNA digested with various restriction enzymes (EcoRI, BamHI, PstI, HindIII, or Kpn I) and probed with clone 2 yielded single bands in each case, suggesting that apo B PI is encoded by a single gene (data not shown).

DISCUSSION

The lack of amino acid sequence information for rat apo B and our consistent inability to obtain heterologous translation of rat apo B mRNA has precluded direct identification of the apo B clones by demonstration of sequence identity or by hybrid-selection/translation, respectively; however, the identity of the clones could be established by immunological characterization of the β-galactosidase–apo B fusion proteins and by hybridization analysis of complementary mRNA.

In regard to the crucial issue of antibody specificity, the antiserum used in the original screens was raised against rat apo B PI, purified as an approximately 400-kDa peptide from

![Fig. 2](image-url)
the modified plasma VLDL (t-VLDL) that accumulates upon intravenous injection of Triton WR1339 (10). The large size of the PI peptide, and the fact that t-VLDL protein is essentially pure apo B (10), argues against contamination of the immunogen with extraneous serum components. Indeed, immunoblotting studies indicated that this antibody reacted specifically with only the apo B polypeptides present in rat plasma or rat liver homogenates, and the antibody showed no reaction with lipoprotein-depleted rat plasma, rat plasma high density lipoproteins (HDL), or purified rat apo E under conditions that gave intense staining of PI and PIII standards as well as of fusion proteins from the positive clones 2 and 49a (Fig. 1). Moreover, preabsorption of the antibody with a large amount of lipoprotein-depleted rat plasma did not affect its reactivity with the fusion proteins or with the complete panel of plaques from positive clones. Conversely, when the antiserum was adsorbed with immobilized rat PI, which had been purified by sequential preparative NaDodSO4/PAGE and high performance gel exclusion chromatography, the reactivity towards the clone 2 and 49a fusion proteins was lost in parallel with that towards the PI and PIII standards. Finally, the affinity-purified antibodies recovered from the PI adsorbant again reacted with both fusion proteins and the apo B standards.

The significantly greater sensitivity of the original anti-PI antiserum for PI as compared to PIII raised the possibility that it would predominantly select clones containing PI-specific apo B sequences. This was indeed the case: all 15 clones remained positive in the plaque assay after the antiserum had been made PI-specific by adsorption with immobilized pure PIII, as did the fusion proteins from clones 2 and 49a on immunoblot analysis. Thus, the four non-overlapping sequences represented in our apo B clones all derive from the region of PI that is absent from PIII. This is consistent with polycistron run-off translation experiments which suggest that the PI-specific epitopes of the PI peptide are located in its COOH-terminal moiety; clearly, a cDNA for the corresponding 3'-terminal end of PI mRNA should be much more abundant in our oligo(dT)-primed library than that of the distant NH2(5')-terminal PI/PIII-common moiety.

Additional strong evidence for the identity of the apo B PI clones comes from the RNA hybridization studies. First, and most important, the cDNA from the two independent clones 2 and 49a both hybridize to the same, very large mRNA species in rat liver, thus providing a direct physical link between these two nonoverlapping sequences, which were originally independently identified by their ability to encode separate PI-specific epitopes. Second, the large size of apo B PI (400 kDa) predicts a very large mRNA with a coding region of about 11 kb. This was indeed observed, since the rat liver poly(A)+ RNA species that hybridizes to the apo B cDNA has a size of about 20 kb (Fig. 3). Third, in vivo pulse-labeling studies indicate that apo B PI accounts for about 0.1% of rat liver total protein synthesis (unpublished data), suggesting that apo B mRNA should be relatively abundant in liver. This is consistent with the relative intensity of the rat liver 20-kb species compared to that for mouse liver apo E, which constitutes about 1% of the mRNA in that tissue. Finally, the tissue distribution of the 20-kb putative PI mRNA is that expected from in vivo pulse labeling studies (unpublished data), which showed abundant synthesis of apo B PIII in both liver and intestine but of apo B PI in liver only, with neither form being synthesized in kidney and spleen. In agreement with this, none of a wider range of extrahepatic tissues showed specific hybridization when probed with PI cDNA, except for rat small intestine.

Unlike apo A-I, apo E, and other apolipoproteins, the sequence for clone 49a shows neither amphipathic helices nor pronounced internal repeats. Although in keeping with the possibility that apo B and apo E may be derived from a common ancestral sequence, since they bind to a common receptor, the significance of the near identity of a 16-base segment with the –10 to +6 base sequence of rat apo E remains to be determined; no other significant homologies at the amino acid or nucleic acid level were observed with other proteins or genes. However, since clone 49a represents only about 2% of the apo B coding region, questions of internal and external homologies will clearly remain until larger segments are sequenced.

The intractability of the apo B peptides has left their basic molecular properties in dispute. We base our size estimates (10) of 400 and 210 kDa for human PI and rat PIII, respectively, on the hydrodynamic properties of the purified peptides in concentrated guanidine hydrochloride. However, previous estimates for human PI obtained by using similar procedures gave the lower value of 250 kDa (4, 6, 7), perhaps...
due to size heterogeneity in the samples used, whereas the higher value of 350 kDa for human PI, which is the basis for the widely used B-100/B-48 centile nomenclature for the apo B peptides, is an apparent molecular mass obtained by preparative NaDodSO₄/PAGE (8). On the other hand, a variety of other groups have reported lower molecular weight polypeptides associated with LDL (30–35) and have suggested that the high molecular weight forms represent aggregates or polymers of these low molecular weight forms, although the origin of the latter as proteolytic artifacts (36) appears more likely. The failure to obtain full-length apo B by translation of liver or intestinal mRNA in heterologous systems has led to further proposals that the mature, large apo B peptides could arise by an unprecedented post-translational ligation of smaller primary translation products (37). This proposal was recently considerably strengthened by reports by Olofsson and co-workers (38, 39) that poly- and monoclonal antibodies to human apo B precipitated an 80-kDa product from an in vitro translation system programmed with human liver RNA. At this point the significance of this very interesting finding is unclear: our group has consistently failed to obtain apo B translation products from liver RNA in heterologous systems (unpublished data), but the results of Janero and Lane (40) on the elongation kinetics of nascent PI peptides in chicken hepatocytes, as well as our results on the run-off translation of nascent PI and PIH peptides from rat liver and intestinal polysomes (unpublished data), are entirely compatible with conventional synthesis of the full-length apo B peptides. Our present finding of the 20-kb mRNA for rat PI certainly argues against its coding for an 80-kDa or smaller peptide, since this would leave an unprecedented 90% noncoding region for this mRNA species.

The resolution to these protracted differences of opinion regarding the structure and biogenesis of the apo B peptides is now at hand through identification of the coding regions for their cloned mRNAs.

Note Added in Proof. We now have additional evidence supporting the identity of the apo B clones. First, immobilized rat fusion proteins bind antibodies that, when eluted, react with apo B PI. Second, human hepatoma cDNA clones, identified by screening of expression libraries with polyclonal antibodies to intact human apo B or to synthetic peptides based upon known human apo B sequences, hybridize to mRNA identical in size to that reported here. Also, recent experiments with intestinal mRNA have revealed higher levels of the 20-kb apo B mRNA, some approaching those present in liver, than those found in the preparations used for the studies reported here.

We are grateful to Camilla Heinzenz, Karen Colinias, and Diana Quon for technical assistance and to Dr. Nancy Young for preparation of rat apo B and characterization of the antibodies. This work was supported in part by grants from the Veterans Administration and from the Public Health Service (HL27481, HL27711, and AM27008) and by Research Grants 760-G1-1 and 764-G1-1 from the Los Angeles Affiliate of the American Heart Association. M.P.Y. is a predoctoral trainee supported by National Research Service Award GM07104. M.M. was supported in part by Atherosclerosis Research Training Grant HL07-386.