ABSTRACT An expression library made in plasmids pUC8 and pUC9 with mRNA derived from the human hepatoma cell line HepG2 was screened with a rabbit antiserum to human low density lipoprotein (LDL). Approximately 12,000 clones were screened and five positives were identified. The cDNA inserts were all 1500-1600 base pairs in length. The insert from one clone, pB8, was isolated, labeled by nick-translation, and found to cross-hybridize strongly with the other four cDNA clones. The pB8 clone produces a fusion protein of ≈37.5 kDa that reacts in electrophoretic transfer blot analysis with rabbit anti-human LDL. This reactivity can be abolished by pretreatment of the antiserum with purified human LDL, p = 1.025-1.050 g/ml. A pB8-derived probe was used to demonstrate that apolipoprotein B (apo B) mRNA is present in HepG2 cells and liver extracts but not in HeLa cells or extracts from small intestine, heart, aorta, spleen, brain, skeletal muscle, lung, kidney, or ovary. RNA transfer blot analysis revealed that HepG2 cell apo B mRNA was ≈22 kilobases in length. These cDNA clones should allow the isolation of the apo B gene and ultimately the elucidation of the primary structure of this protein.

Apolipoprotein B (apo B) is abundant in plasma, with a concentration of 0.7-1.0 mg/ml, and is found as a constituent of chyomicrons, very low density lipoproteins, and low density lipoproteins (LDL) (1). Most plasma apo B resides in LDL, a particle containing 75% lipid and 25% protein, almost all of which is apo B. LDL cholesterol levels are correlated with coronary disease susceptibility and recently the same association has been found for apo B levels (2). Human apo B is a glycoprotein that can be separated into two forms on NaDodSO4/PAGE, designated B-100 and B-48 (3). B-100 is made in liver, a necessary step for the synthesis and secretion of hepatic-derived, triglyceride-rich lipoproteins, and appears to have a molecular mass of ≈550 kDa (3, 4). B-48 is synthesized in the small intestine, plays an important role in synthesis and secretion of intestinal-derived, triglyceride-rich lipoproteins, and appears to have a molecular mass of ≈275 kDa (3, 4).

The current report is concerned with the isolation of cDNA clones for human apo B. Recently, cDNA clones have been isolated for the other human apolipoproteins (5-11). In most cases, the method used relied on previously reported amino acid sequence data obtained by protein chemistry techniques. Sequences of amino acids specified by relatively unambiguous codons were identified and the corresponding mixture of oligonucleotides was synthesized. These were used as probes to screen liver cDNA libraries and appropriate clones were selected. Unfortunately, after delipidation of LDL, apo B is quite insoluble (12) and has resisted characterization by standard protein chemistry techniques. As a result, very little apo B primary amino acid sequence is known (13) and a different approach had to be taken for cloning apo B cDNA. Therefore, in the current study, apo B cDNA clones were isolated from an expression library made from human hepatoma cell mRNA by utilizing a polyclonal antiserum to human LDL.

METHODS

Construction of HepG2 cDNA Expression Library. An expression library was prepared by using the plasmid vectors pUC8 and pUC9. Vectors were digested to completion with Pst I and EcoRI, isolated by gel filtration, and ligated to cDNA made in the following manner. Total RNA was prepared from human hepatoma cell line HepG2. Poly(A)+ mRNA was isolated by adsorption to and elution from oligo(dT)-cellulose. The first strand of the DNA was synthesized with avian myeloblastosis virus reverse transcriptase (BRL) by using oligo(dT)12-18 primers. After boiling to dissociate the RNA-DNA hybrids, the second strand was synthesized by using the Klenow fragment of DNA polymerase I (BRL). Pst I and EcoRI linkers were added sequentially to the double-stranded cDNA as described by Helfman (14). For the pUC8 vector, Pst I linkers were attached to the end of the cDNA corresponding to the 3' end of the poly(A)+ mRNA and EcoRI linkers were attached to the 5' end. The orientation of the linkers was reversed for the pUC9 vector. The appropriate vector and cDNA with linkers were ligated and transformed into Escherichia coli strain DH1. Transformants were selected for the acquisition of the plasmid ampicillin-resistance gene.

Immunoscreening of the HepG2 cDNA Expression Library. Rabbit anti-human LDL (i.e., anti-human apo B) antiserum was a gift from D. Williams (State University of New York, Stony Brook). This antiserum reacts specifically with apo B and with no other human apolipoproteins on electrophoretic transfer blot analysis. The bacterial colonies were grown on LB plates containing ampicillin and transferred with a toothpick to nitrocellulose filters. The nitrocellulose filters were placed on fresh ampicillin plates for further growth. The bacteria were then lysed and screened with anti-human LDL antiserum according to the method of Helfman (14). After removal of antiserum and extensive washing with Tris-buffered saline (50 mM Tris/HCl, pH 7.5/150 mM NaCl), the filters were incubated with 5 × 10^6 cpm of 125I-labeled protein A (125I-protein A) (Amersham; specific activity ≈ 30 mCi/ mg; 1 Ci = 37 GBq) for 1 h at room temperature. The filters were then washed extensively with Tris-buffered saline, air dried, and exposed to x-ray film.

Preparation of Plasmid DNA and Radioactive cDNA Probe. Bacterial clones were grown in Luria broth supplemented with antibiotics and chloramphenicol. The plasmid DNA was prepared as described by Maniatis et al. (15). The cDNA probe was prepared by nick translation using bacteriophage T4 DNA polymerase and [α-32P]ATP (16). The radioactive cDNA probe was precipitated with ethanol and were resuspended in 0.2 M NaCl.

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: apo B, apolipoprotein B; apo A-I, apolipoprotein A-I; LDL, low density lipoprotein(s); kb, kilobase(s).
with ampicillin (50 μg/ml). Plasmid DNA was isolated by using the alkaline lysis method and was further purified on a CsCl gradient (15).

Plasmid DNA from clone pB8 was digested with Pst I and EcoRI (New England Biolabs) to obtain the cDNA insert, which was isolated on a 1% low-melting agarose gel, radio-labeled with [α-32P]dATP by nick-translation, and used as a probe for DNA or RNA hybridizations.

**DNA Dot Blot Analysis**. Plasmid DNA (200 ng in 10 μl) from pB2, -3, -4, -5, or -8 or from pUC8 (control) was heated to 100°C for 10 min and chilled on ice. The DNA sample was then denatured with 10 μl of 1 M NaOH at room temperature for 20 min and neutralized with 10 μl of 0.5 M Tris-HCl (pH 8.0) buffer containing 1 M NaCl/0.3 M sodium citrate/1 M HCl as described (15). The samples were then chilled on ice, dot blotted onto nitrocellulose filters, air dried, and baked at 80°C for 2 hr. The filter was hybridized with 32P-labeled pB8 at 65°C in a buffer containing 0.15 M sodium citrate, 1.5 M NaCl, 0.2% Ficoll, 0.2% polyvinylpyrrolidone, 0.2% bovine serum albumin, and salmon sperm DNA at 50 μg/ml. After washing with 15 mM NaCl/1.5 mM sodium citrate for 15 min, the filter was exposed to x-ray film.

**Electrophoretic Transfer Blot Analysis**. Overnight cultures of bacterial clones were pelleted by centrifugation, resuspended with 0.1 vol of Laemmli sample buffer, and boiled for 2 min (16). The bacterial lysates were subjected to NaDodSO4/PAGE in 10% acrylamide (16). Fractionated proteins either were stained with 0.1% Coomassie brilliant blue or electrotransferred to nitrocellulose filters as described by Towbin (17). After blocking, the filters were incubated for 1 hr with either nonimmune rabbit serum, rabbit anti-human LDL, or rabbit anti-human LDL adsorbed with LDL. In the latter case, LDL (μ = 1.025-1.050 g/ml) was isolated by ultracentrifugation and the antiserum was preincubated with excess LDL for 1 hr at 37°C to adsorb the specific reactivity to LDL. The filters were then washed with Tris-buffered saline, probed with 125I-protein A, washed again, and exposed to x-ray film.

**RNA Dot Blot Analysis**. Total RNA from human cell lines or rabbit tissues was isolated by the guanidinium isothiocyanate/CsCl method (18). Formaldehyde was added to a final concentration of 3% to RNA samples that were then heated at 60°C for 15 min. The solutions were then diluted to a final volume of 100 μl with 3 M NaCl/0.3 M sodium citrate to give the desired RNA concentrations. These were dot blotted onto nitrocellulose filters with a template manifold apparatus (Schleicher & Schuell). The filters were air dried, baked at 80°C for 90 min, and hybridized to 32P-labeled pB8 cDNA probe as described for DNA dot blot analysis. The filters containing human or rabbit RNA were hybridized and washed at 65°C and 55°C, respectively.

**RNA Transfer Blot Analysis**. Total HepG2 RNA (20 μg) was denatured by heating to 60°C in 6% formaldehyde/50% formamide and then fractionated by electrophoresis in a 0.6% agarose/6% formaldehyde gel. The fractionated RNA was transferred to a nitrocellulose filter and hybridized to 32P-labeled pB8 cDNA as in the DNA dot blot analysis. As a control, the filter was also hybridized in an identical manner to an apolipoprotein A-I (apo A-I) cDNA probe derived from a previously isolated clone, pAI-121 (19).

**RESULTS**

**Identification of apo B cDNA Clones**. About 12,000 cDNA clones constructed from HepG2 poly(A)+ RNA were digested on nitrocellulose filters and screened with rabbit anti-human LDL antiserum. Six apparently positive clones from the first screening were streaked out to obtain single colonies and then were rescreened with the antiserum (Fig. 1). Clones pB2, -3, -4, -5, and -8 reacted strongly with the antiserum (Fig. 1, bottom), whereas clones 6 and the control cDNA clone, which expresses a β-galactosidase–human antithrombin III fusion protein.

**Cross-Hybridization of apo B cDNA Clones**. Plasmid DNAs prepared from clones pB2, -3, -4, -5, and -8 were dot blotted onto nitrocellulose filters and hybridized to the 32P-labeled pB8 cDNA probe. pB8 cDNA hybridized strongly with the other four clones but not with control pUC8 plasmid DNA (Fig. 2) or with the clone that expresses the β-galactosidase–human antithrombin III fusion protein (data not shown). This indicates that the cDNA sequence of pB8 overlaps that of the pB2, -3, -4, and -5 clones. This was verified by preliminary DNA sequence analysis (data not shown).

**Identification of a Fusion Protein from the pB8 cDNA Clone**. Bacterial extracts from clone pB8 and a control clone containing the plasmid pUC8 without a cDNA insert were compared by NaDodSO4/PAGE. After Coomassie brilliant blue staining, the pB2, -3, -4, -5, and -8 bands were resolved, whereas the control clone was nonreactive and clone C is a control that expresses a β-galactosidase–human antithrombin III fusion protein.

**RESULTS**

**Identification of apo B cDNA Clones**. About 12,000 cDNA clones constructed from HepG2 poly(A)+ RNA were digested on nitrocellulose filters and screened with rabbit anti-human LDL antiserum. Six apparently positive clones from the first screening were streaked out to obtain single colonies and then were rescreened with the antiserum (Fig. 1). Clones...

**Cross-Hybridization of apo B cDNA Clones**. Plasmid DNAs prepared from clones pB2, -3, -4, -5, and -8 were dot blotted onto nitrocellulose filters and hybridized to the 32P-labeled pB8 cDNA probe. pB8 cDNA hybridized strongly with the other four clones but not with control pUC8 plasmid DNA (Fig. 2) or with the clone that expresses the β-galactosidase–human antithrombin III fusion protein (data not shown). This indicates that the cDNA sequence of pB8 overlaps that of the pB2, -3, -4, and -5 clones. This was verified by preliminary DNA sequence analysis (data not shown).

**Identification of a Fusion Protein from the pB8 cDNA Clone**. Bacterial extracts from clone pB8 and a control clone containing the plasmid pUC8 without a cDNA insert were compared by NaDodSO4/PAGE. After Coomassie brilliant blue staining, the pB2, -3, -4, -5, and -8 bands were resolved, whereas the control clone was nonreactive and clone C is a control that expresses a β-galactosidase–human antithrombin III fusion protein.
blue staining, a protein of 37.5 kDa was observed in the extract of the pB8 clone but not in the control (Fig. 3A). This protein reacted with rabbit anti-human LDL antiserum (Fig. 3C) but not with nonimmune rabbit serum (Fig. 3B) or rabbit anti-human LDL antiserum preadsorbed with LDL (Fig. 3D). In no case did bacterial proteins from the control clone react (Fig. 3 B–D). These results indicate that the pB8 cDNA clone produces a fusion protein that is specifically recognized by an anti-human LDL (apo B) antiserum.

**Tissue Distribution of apo B mRNA.** Total RNAs prepared from human cell lines and rabbit tissues were dot blotted onto nitrocellulose filters and hybridized to the 32P-labeled pB8 cDNA probe. Fig. 4A shows that apo B mRNA is expressed in a human hepatoma cell line, HepG2, but not in undifferentiated HeLa cells. Fig. 4B shows that the pB8 cDNA probe hybridizes, under less stringent conditions, with rabbit liver RNA but not rabbit RNA from small intestine, heart, aorta, spleen, brain, skeletal muscle, lung, kidney, and ovary.

**Size of apo B mRNA Expressed in HepG2 Cells.** Total RNA isolated from HepG2 cells was analyzed in RNA transfer blot analysis by utilizing the pB8 cDNA probe. In several RNA preparations, including a poly(A)+ mRNA preparation, a major band of \( \approx 22 \) kb was detected (Fig. 5). Beneath this band, a smear was seen that varied in intensity in different RNA preparations. In the same RNA preparations, an apo A-I probe gave a single band of the expected size (0.95 kb). Therefore, the RNA preparations are probably adequate and the smear seen with the apo B probe is most likely due to difficulties related to the isolation of very large RNAs.

![Fig. 3](image-url)  
**Fig. 3.** Electrophoretic transfer blot analysis of the pB8 clone fusion protein. Bacterial proteins from overnight cultures were prepared, separated by NaDodSO4/PAGE on 10% polyacrylamide gels, and electrotransferred onto a nitrocellulose filter. In A–D, lanes 1 contain bacterial proteins from E. coli strain DH1 transformed with pUC8 (no cDNA insert), and lanes 2 contain bacterial proteins from clone pB8. Molecular mass standards are shown in kDa. (A) Gel stained with 0.1% Coomassie brilliant blue. The arrow indicates the fusion protein of 37.5 kDa produced by the pB8 clone. (B–D) Autoradiograms of filters incubated for 1 hr with the indicated antiserum, washed, probed with 125I-protein A, and exposed to x-ray film for 5 hr. Filters from electrophoretic transfer blots were incubated with nonimmune rabbit serum (B), rabbit anti-human LDL (C), and rabbit anti-human LDL, preadsorbed with LDL (\( p = 1.025-1.050 \) g/ml) (D).

![Fig. 4](image-url)  
**Fig. 4.** Tissue distribution of apo B mRNA. Total RNA isolated from human HepG2 cells, HeLa cells, and 10 rabbit tissues was dot blotted onto nitrocellulose filters. Filters were hybridized to 32P-labeled pB8 cDNA probe, washed, and exposed to x-ray film for 24 hr. In A, hybridization and washing were at 65°C. In B, hybridization and washing were at 55°C.

**DISCUSSION**

A cDNA expression library, made from mRNA of a human hepatoma cell line, HepG2, has been used to isolate five cross-hybridizing clones. These clones produce a fusion protein immunoreactive with polyclonal antisera to human LDL mRNA hybridizing with these clones is found only in liver and not in intestine or other tissues. Hepatic mRNA is \( \approx 22 \) kb in length.

Analysis of plasma apo B by NaDodSO4/PAGE reveals two principal forms. The larger is called B-100 and the smaller is designated B-48 (4). In humans, B-100 and B-48 have apparent molecular masses of 550 kDa and 275 kDa, respectively (3). In most species, including man, B-100 is made in liver, whereas B-48 is made in intestine (4, 20–24). In the rat, B-100 and B-48 are made in liver, but only B-48 is made in intestine. The relationship between B-100 and B-48 has been explored with monoclonal antibodies (25, 26). With B-100 as the immunogen, several sets of monoclonal antibodies have been developed. B-100 epitopes have been mapped in a linear nonrepetitive array. For a subset of these epitopes, the monoclonal antibodies disrupt apo B binding to the LDL receptor but do not bind B-48. For another subset of epitopes, the monoclonal antibodies bind B-48 but do not disrupt receptor binding (25). Thus, B-48 and B-100 are antigenically
Biochemistry: Huang et al.

**FIG. 5.** RNA transfer blot analysis of total RNA from HepG2 cells with apo B and apo A-I cDNA probes. Total RNA from the HepG2 cell line was separated on a 0.6% agarose/6% formaldehyde gel and blotted onto a nitrocellulose filter. The filter was hybridized to 32P-labeled pB8 cDNA (left lane) or pAI-121 cDNA (right lane). The graph indicates sizes of the hybridizing mRNA species. The standard curve (solid line) was derived by using 28S rRNA (5 kb) and 18S rRNA (2 kb) as standards. The dashed portion of the line is the extrapolation to the migration distances of apo B and apo A-I mRNA.

The sizes of these mRNA species are indicated in kb on the ordinate.

related, and it has been suggested that B-48 represents either the amino- or carboxyl-terminal half of the B-100 protein and that this part of apo B is not involved in receptor binding (25).

The current study may reveal information about the relationship between these two forms of apo B. The cDNA library we used was constructed by priming first-strand synthesis from the poly(A) tail of HepG2 mRNA. Therefore, the clones isolated should all contain sequence corresponding to the 3' untranslated region of apo B mRNA (confirmed by preliminary sequence analysis) and the coding region for the carboxyl-terminal amino acids of B-100. The length of the cDNA clones, 1500–1600 bp, and the molecular mass of the fusion protein produced, 37.5 kDa, suggest that the 3' untranslated region of apo B-100 mRNA is ~500 base pairs and the clones contain the codons for the carboxyl-terminal 300–350 amino acids of B-100. The fact that these clones fail to hybridize to intestinal mRNA together with the results from the monoclonal antibody studies suggest that B-48 corresponds to the amino- but not the carboxyl-terminal region of B-100.

Due to its difficult physicochemical properties, there is a paucity of data on apo B structure. Currently, there are two theories. One states that B-100 is a single polypeptide chain with an apparent molecular mass of 550 kDa on NaDodSO4/PAGE (3). If this is so, B-100 would be one of the longest polypeptides in nature. The cited studies above with monoclonal antibodies, suggesting that B-100 epitopes map in a linear nonrepetitive array, support this theory. The other theory states that B-100 is composed of multiple subunits of small monomers (27–30). A very recent study has provided evidence for a monomeric molecular mass of 19 kDa (31). If the first theory is correct, B-100 mRNA must contain coding information for ~5000 amino acids and therefore be at least 15 kb in length. If the second theory is correct, B-100 mRNA would need to code for only ~200 amino acids and could be quite small. The current study may provide insight into B-100 structure by showing that apo B mRNA is ~22 kb in size. mRNA of this length could contain enough coding information for >5000 amino acids and could specify a 550-kDa protein. It would be improbable to have an mRNA of this size coding for a 19-kDa protein. Thus, the evidence from the current study favors the theory that apo B is a single, long, polypeptide chain. However, it is still possible that apo B mRNA codes for a large protein precursor, which is processed to a small subunit. Until more is known about the mRNA sequence, judgment as to which theory is correct must be withheld.

Our isolation of a B-100 cDNA clone is a first step toward understanding apo B mRNA structure and deducing the complete primary amino acid sequence of apo B. As noted, apo B plays a central role in lipoprotein metabolism, and human variation in apo B gene structure or the regulation of apo B gene synthesis is thought to be a cause of human dyslipoproteinemias or atherosclerosis susceptibility (or both) (32, 33). Knowledge of normal apo B gene structure will allow comparisons with the structure of apo B genes isolated from patients to verify whether this is in fact the case.

We thank Dr. B. Knowles for providing the HepG2 cells used to prepare the RNA used in library construction. This research was supported by National Institutes of Health Grants HL33714 and HL32435 (J.L.B.) and HL30712 (S.C.B.) and by American Heart Association Grant 83-1202 (S.C.B.). J.L.B. is an Established Investigator of the American Heart Association.


