Identification, Characterization, Cloning, and Expression of Apolipoprotein C-IV, a Novel Sialoglycoprotein of Rabbit Plasma Lipoproteins*

Lin-Hua Zhang, Leila Kotite, and Richard J. Havel†
From the Cardiovascular Research Institute and Department of Medicine, University of California, San Francisco, California 94143-0130

We have identified and characterized a novel proline- and arginine-rich protein component of lipoproteins, present in up to five sialylated isoforms, in rabbit blood plasma. The pl of the desialylated protein is 5.7. Based upon its N-terminal sequence, a complete cDNA sequence of 555 nucleotides was cloned from rabbit liver. The synthesized protein is predicted to contain 124 amino acids, including a typical signal peptide of 27 residues. The mature protein of 97 amino acids, designated apolipoprotein C-IV, is associated with the lipoproteins of blood plasma, primarily very low density and high density lipoproteins. It contains two potential amphipathic helices characteristic of plasma apolipoproteins and forms discoidal micelles with phosphatidylcholine. Northern analysis shows a single 0.6-kb apolipoprotein C-IV mRNA, detected only in the liver, and Southern analysis suggests a single copy gene. Sialylated apolipoprotein C-IV is secreted from transfected mammalian cells. Nucleotide sequence comparisons demonstrate a strong homology to portions of the upstream regions of the mouse and human apolipoprotein C2 genes, within each of which a distinct gene has recently been identified. The nucleotide sequences and the predicted amino acid sequences, as well as corresponding cDNA sequences in the rat and monkey, indicate that the apolipoprotein C4 gene has been highly conserved during mammalian evolution.

Lipoproteins in blood plasma of mammals contain a variety of protein components, most of which are water-soluble proteins, encoded by a gene family characterized by repeating sequences of amphipathic helices containing 11 or 22 amino acids (1). Those of low molecular mass (approximately 6–12 kDa) are generally called "C" apoproteins, of which three are well recognized (2). These apoproteins are found in triglyceride-rich lipoproteins (chylomicrons and very low density lipoproteins (VLDL)) and high density lipoproteins (HDL), among which they readily transfer according to the relative concentrations and, presumably, surface areas of the lipoprotein particles (3). ApoC-I and apoC-II lack carbohydrate, whereas apoC-III has an O-linked carbohydrate moiety containing up to three sialic acid residues (2). The functions of the C apoproteins are only partly understood. During postprandial lipemia, C apoproteins move from HDL to nascent chylomicrons and are returned to HDL as chylomicrons are metabolized by lipoprotein lipase (4). ApoC-II is a specific cofactor for the action of lipoprotein lipase upon the ester bond of triacylglycerols and glycerophosphatides of triglyceride-rich lipoproteins (5,6). The other C apoproteins can inhibit this action of apoC-II (5,6), and all of the C apoproteins can inhibit the clearance of remnants of triglyceride-rich lipoproteins by the liver (7).

The C apoproteins are most commonly visualized in triglyceride-rich lipoproteins by isoelectric focusing electrophoresis in polyacrylamide gels. During studies of lipoprotein metabolism in the rabbit, we observed a complex pattern of apoproteins in triglyceride-rich lipoproteins (8), unlike that found in other mammals. This pattern, which was also seen in the C apoproteins separated from other soluble apoproteins by gel permeation chromatography, was found to reflect the presence of a second sialoglycoprotein, in addition to apoC-III. Based upon the N-terminal amino acid sequence, which was unusually rich in proline, we have now cloned the cDNA and expressed this protein which was found to be homologous to that of a recently identified gene in the apoC3-apoC2 gene cluster on mouse chromosome 7 (9).

MATERIALS AND METHODS
Preparation of Lipoproteins and Apolipoproteins—Blood was obtained from ear veins of unanesthetized rabbits or by aortic puncture of rabbits anesthetized with ketamine/xylazine that had been fed normal chow or chow enriched with 1% cholesterol. Two rabbits fed normal chow were made diabetic 3–5 days before blood was obtained by intravenous injection of alloxan (200 mg/kg) (10). To assess the acute-phase response of apoC apoproteins, rabbits were bled 48 h after subcutaneous injection of 0.5 ml of turpentine per kg of body weight (11) or an equal volume of 0.15 M NaCl. Blood was collected into tubes containing disodium EDTA, benzamidine, and phenylmethylsulfonyl fluoride to give final concentrations of 0.5, 0.3, and 0.1 mg/ml, respectively. Plasma was obtained after centrifugation at 2000 rpm for 20 min and subjected to ultracentrifugation at densities of 1.006, 1.063, and 1.21 g/ml to obtain lipoprotein fractions: VLDL, intermediate density lipoproteins and low density lipoproteins (IDL + LDL), and HDL (12). For purification, lipoprotein fractions were recentrifuged at the upper density limit. To remove sialic residues from apolipoproteins, lipoproteins were incubated with 0.1 unit neuraminidase from Clostridium perfringens (Boehringer Mannheim) per mg of protein in 0.02 M sodium acetate, pH 5.4, for 2 h at 37 °C.

Apolipoproteins were obtained after delipidation of lipoprotein fractions with ethanol:ether (3:1) (13). C apoproteins were separated from apolipoprotein C3 by gel filtration on Sephadex G-200 (Pharmacia Biotech Inc.) (14). Individual apolipoproteins were separated by preparative isoelectric focusing electrophoresis as described (15) or by anion exchange chromatography on...
Apolipoprotein C-IV, a New Component of Rabbit Lipoproteins

1777

and lipoprotein fraction was determined by enzymatic procedures (23). The content of total cholesterol in plasma after the initial injection. To remove trace immunoreactivity against apoE, the antiserum was recycled through a rabbit apoE affinity column (22). The purified antiserum showed no detectable immunoreactivity against rabbit apoE, apoC-I, or apoC-III. Content of total cholesterol in plasma was determined by enzymatic procedures (23).

Other Analytical Procedures—Western blotting was carried out after transfer of proteins from polyacrylamide gels to nitrocellulose paper (24) with diluted antisera against apoC-IV followed by 125I-labeled goat anti-guinea pig IgG (Sigma) and autoradiography. Unfamilial vesicles of dimyristoylphosphatidylcholine, before and after incubation with apoproteins, were visualized by electron microscopy after negative staining (25). The cofactor property of apoproteins for the activity of lipoprotein lipase from bovine milk was assayed as described (6).

cDNA Cloning—A reverse transcription-polymerase chain reaction (RT-PCR)-based cDNA cloning strategy (26) was used to isolate a cDNA encoding for apoC-IV. Total RNA was isolated from rabbit liver by the method of Chirgwin et al. (27) and桑本友也 et al. (28). Single-stranded cDNA was synthesized with 250 units of SuperScript II RNase H- (Life Technologies, Inc.) from 3 μg total RNA after priming with 30 pmol of a cDNA synthesis primer (CDS primer, 5'CCCCGAATTCCTTTTTTTTTTTTTTTTTT(A/G)(A/G)(T/C)3'); each RT product was then diluted 500-fold and used in subsequent PCR reactions.

Based on the N-terminal amino acid sequence of apoC-IV: EEPEGP[n(7)]PLP[ga(ga(ga(ESRSW)I)](IV) (paraproteins of homologous and lower case letters indicate weak signals), degenerate primers were synthesized (Operon Co.) complementary to the first 9 amino acid residues: P153, 5'G(A/G)CCGA(A/G)GGAGNCNNCAAT(T/C)CCNC(CT/T)3' (N = A, G, C, or T); and the nested primer: P252, 5'CICIC(C/T)TICCIC(C/G)CAG(A/G)AG, T = deoxyinosine) to amino acid residues indicated in Fig. 2B. The cDNA was first screened by three successive rounds of PCR. The first round was asymmetric, with CDS-primed single-stranded cDNA as the template. Amplification was carried out with P153 (4 pmol) as primer in 67 μl Tris (pH 8.3), 12 μM (NH4)2SO4, 2 mM MgCl2, 17 μM bovine serum albumin, 0.15 mM concentration of each of dATP, dGTP, dCTP, and dUTP, and 50 μU of Taq polymerase (Life Technologies, Inc.) in a thermal cycle (Perkin-Elmer) for 30 cycles at 94°C, 45 s; 50°C, 60 s; 72°C, 1 min, with a last extension at 72°C for 7 min. One μl of the first round PCR product was used in the second round PCR amplification, by 8 pmol each of CDS and P153 under the same conditions for 30 cycles. Two bands (major, 380 bp; minor, 330 bp) were resolved by electrophoresis in a 1.5% agarose gel containing 10 μM sodium sulfate, 0.1% SDS for 10 min. The final wash was carried out in 0.1% SSC at 50°C for 15 min and was followed by autoradiography.

Expression and Secretion from Mammalian Cells—Chinese hamster ovary (CHO)-K-1 cells were grown in monolayer in 10-cm dishes containing 10 ml of Dulbecco's modified Eagle's medium-Ham's 21 (DME-H21), 10% fetal calf serum, containing 1.1 mg of sodium pyruvate, and nonessential amino acids in a humidified atmosphere of 5% CO2, 95% air. COS-7 African green monkey kidney cells were grown in DME-H21 medium containing 4.5 g of glucofector and 10% FBS as described (35, 36). A full-length apoC-IV cDNA lacking the poly(A) tail was subcloned into the EcoRI sites of the expression vectors pET3c containing a 14-aminooxy linker sequence. Another construct was made by fusion of the Ndd-BamHI fragment of apoC-IV cDNA encoding only the 97-amino acid mature protein into vector pET3b. Escherichia coli BL21(DE3) cells, transformed with these plasmids, were then induced with 0.4 mM isopropyl-1-thio-β-D-galactopyranoside to an optical density of 0.5 at 600 nm for 2–4 h (34). The whole cell extracts were resolved on SDS-15% polyacrylamide gel, transferred to nitrocellulose membranes, and subjected to autoradiography as described above except that the final wash steps were done twice at 65°C for 15 min. Expression in E. coli—A BamHI-BamHI fragment of apoC-IV cDNA coding the 124-amino acid polypeptide was produced by PCR and inserted into the BamHI fusion site of the expression vector pET3c containing a 14-aminooxy linker sequence. Another construct was made by fusion of the Ndd-BamHI fragment of apoC-IV cDNA encodinonly the 97-amino acid mature protein into vector pET3b. Escherichia coli BL21(DE3) cells, transformed with these plasmids, were then induced with 0.4 mM isopropyl-1-thio-β-D-galactopyranoside and a 10% FBS as described (35, 36). Two μl of plasmid prepared by density ultracentrifugation in cesium chloride (28) or a modified-large-scale alkaline lysis (39) followed by transformation of competent E. coli cells with RNase K and each of the above 2 constructs resulted in transformation efficiency of 10 106 transfectants/μg DNA. The transformed cells were grown in 2 ml of s-32 g L-lysine, 3 μM potassium ferrocyanide, 1 μM MgCl2, and 0.5 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside (36) or by β-galactosidase assay as described (35). After a 24-h incubation, the tranducted cells were removed and cells were rinsed with 5 ml of 10% dimethyl sulfoxide in phosphate-buffered saline. Ten ml of fresh medium were added, and incubation was continued for 48 h before collection of the medium.

For immunoprecipitation, 2 μl of anti-apoC-IV antiserum was added to 5 ml of the medium collected from each dish of transfected cells and the mixture, supplemented with 2 μg/ml each of apoprotein, leupeptin, and pepstatin A, and 0.02% sodium azide, was incubated with rotation at 4°C overnight. Then, 80 μl (0.1 g/ml) of protein A-linked Sepharose CL-4B beads (Pharmacia) were added and incubation was continued for
RESULTS

Isolation and Characterization of ApoC-IV—Apolipoprotein C-IV was found during a systematic analysis of small molecular weight apolipoproteins (C apolipoproteins) of very low density lipoproteins (VLDL) from rabbit blood plasma. Upon isoelectric focusing, electrophoresis of apoVLDL from normal rabbits, several components were observed with isoelectric points more acidic than apoE were evident (Fig. 1, lane 3). These proteins were also seen upon isoelectric focusing electrophoresis of the apoC apolipoprotein fraction separated from apoVLDL by gel filtration chromatography (Fig. 2, lane 1). The amino acid composition of the three most acidic of these proteins, isolated by preparative isoelectric focusing (Fig. 2), resembled that of apoC-III from other species, except that the component with a pI of 4.6 (designated "apoC-IV") contained somewhat more proline and arginine than apoC-III (Fig. 2, lane 3). The component with a pI of 5.0 (third lane) had cofactor properties for lipoprotein lipase from bovine milk (data not shown), but its amino acid composition differed substantially from that expected for apoC-III. The amino acid composition of the components with pI 5.2 and 4.8 (second and fourth lanes) was similar, and they were particularly rich in proline and arginine (designated "apoC-IV-L") contained somewhat more proline and arginine. The composition of the other components differed from that of known C apolipoproteins in other species. The component with a pI of 5.0 (third lane) had cofactor properties for lipoprotein lipase from bovine milk (data not shown), but its amino acid composition differed substantially from that expected for apoC-III.

The amino acid composition of the components with pI 5.2 and 4.8 (second and fourth lanes) was similar, and they were particularly rich in proline and arginine (designated "apoC-IV-L") contained somewhat more proline and arginine. The composition of the other components differed from that of known C apolipoproteins in other species. The component with a pI of 5.0 (third lane) had cofactor properties for lipoprotein lipase from bovine milk (data not shown), but its amino acid composition differed substantially from that expected for apoC-III.

The individual proteins were also isolated from the C apolipoprotein fraction by ion exchange chromatography (Table I). In addition to the three proteins separated by preparative isoelectric focusing, a major component was found with an amino acid composition that resembled human apoC-I together with a component tentatively identified as the 12-kDa thrombolytic fragment of apoE (42). The novel proline- and arginine-rich protein (PARP) thus appeared to be a sialoglycoprotein, normally present as isoforms with up to 4 or 5 sialic acid residues. Its apparent molecular mass on SDS-gel electrophoresis was about 14 kDa. This was confirmed by two-dimensional electrophoresis with Coomassie Blue staining and immunoblotting (Fig. 3), which also showed that the apparent molecular mass of the sialylated forms of PARP was greater than that of the desialylated protein. PARP contained approximately 2 cysteine residues (Table I). Upon immunoblotting apoVLDL separated by SDS gel electrophoresis without addition of reducing agents, bands of higher molecular weight, consistent with homodimers and homotrimers, were observed (data not shown). These components were absent in the presence of reducing agents, consistent with disulfide linkage of the oligomers.

PARP was also evident upon isoelectric focusing of apoHDL (data not shown). By radioimmunoassay, PARP was distributed mainly between VLDL and HDL of a rabbit fed regular chow, with undetectable amounts in IDL, LDL, and lipoprotein-free serum (α > 2.1 g/ml), whereas PARP was detected only in VLDL from a cholesterol-fed rabbit (Table II). PARP was also found in apoVLDL isolated from perfusates of isolated rabbit livers (data not shown). Thus, PARP appeared to be a

| Amino acid | ApoC-L | ApoE fragment | PARP (apoC-IV) | ApoC-II | ApoC-III |
|------------|--------|---------------|----------------|---------|----------|
| Lysine     | 13.0   | 5.7           | 3.2            | 7.2     | 5.2      |
| Histidine  | 1.8    | 1.4           | 1.2            | 2.7     | 2.5      |
| Arginine   | 3.9    | 10.3          | 9.9            | 2.3     | 1.4      |
| Asparagine | 6.7    | 6.7           | 9.2            | 11.6    | 8.3      |
| Threonine  | 10.9   | 3.2           | 6.8            | 9.8     | 5.1      |
| Serine     | 8.6    | 4.3           | 5.8            | 9.6     | 12.3     |
| Glutamic acid | 16.0 | 20.4          | 13.4           | 9.7     | 15.2     |
| Proline    | 3.3    | 4.6           | 10.2           | 4.8     | 4.2      |
| Glycine    | 2.8    | 5.9           | 8.9            | 4.3     | 7.1      |
| Alanine    | 4.9    | 15.2          | 7.1            | 13.9    | 16.0     |
| Cysteine   | NA     | NA            | NA             | NA      | NA       |
| Valine     | 2.0    | 6.8           | 4.7            | 4.6     | 5.1      |
| Methionine | 0.9    | 0.5           | NA             | NA      | 16       |
| Isoleucine | 8.9    | 2.2           | 0.6            | 2.6     | 1.3      |
| Leucine    | 7.7    | 8.7           | 14.7           | 12.6    | 7.8      |
| Tyrosine   | 0.0    | 0.9           | 2.4            | 5.8     | 2.6      |
| Phenylalanine | 8.6  | 3.2           | 2.4            | 0.5     | 4.3      |
| Tryptophan | NA     | NA            | NA             | NA      | NA       |

* Presumed to be 12-kDa thrombolytic fragment of apoE (42).

NA, not analyzed. Approximately 2 mol/100 mol of Cys, assayed as cysteic acid (20), were found in another preparation of PARP.
Apolipoprotein C-IV, a New Component of Rabbit Lipoproteins

FIG. 3. Two-dimensional gel electrophoretic analysis of apoVLDL (A and C) and desialylated apoVLDL (B and D). Gels from one preparation of apoVLDL were stained with Coomassie Blue R-250 (A and B) and those from another preparation of apoVLDL were subjected to Western blotting with apoC-IV antiserum (C and D). Cathode was at left and anode at right of each gel.

Distribution of PARP (apoC-IV) in ultracentrifugally separated lipoproteins in rabbit blood plasma

The concentration of apoC-IV in blood plasma from normal and cholesterol-fed rabbits, as estimated by radioimmunoassay with purified desialylated apoC-IV, was in the range of 0.5–1.2 mg/dl. The immunoreactivity of sialylated forms of apoC-IV has not been determined and the actual concentrations of apoC-IV may exceed these values.

Table II

| Lipoprotein fraction | % concentration in whole plasma |
|----------------------|---------------------------------|
| VLDL                 | 40                              |
| IDL                  | ND                              |
| LDL                  | ND                              |
| HDL                  | 48                              |
| Lipoprotein-poor plasma | ND                         |

a Plasma total cholesterol = 60 mg/dl.
b Plasma total cholesterol = 335 mg/dl.
c ND, not detected.

FIG. 4. Negatively stained preparations of discoidal complexes of apoC-IV with dimyristoyl phosphatidylcholine with apoC-IV (top) and rabbit apoE (bottom). For each preparation, unilamellar liposomes of dimyristoylphosphatidylcholine prepared in a French pressure cell (120–170 μg) were mixed with one-fifth the mass of apoC-IV or apoE in 0.25 ml of 0.15 M NaCl and incubated at 37 °C for 1 h before negative staining. × 180,000.

b bona fide apolipoprotein that was distributed in a manner resembling that of other C apoproteins. Based upon these properties, PARP was provisionally designated as apoC-IV.

Properties of ApoC-IV—Since apoC-IV was first isolated from plasma of alloxan diabetic rabbits, in which increased amounts of serum amyloid A were also expressed (see also Fig. 2, first lane), we considered the possibility that apoC-IV is an acute phase protein. Within 48 h after administration of turpentine subcutaneously to two normal rabbits, the apparent mass of serum amyloid A was markedly increased, whereas that of apoC-IV was not (Fig. 1); furthermore, the concentration of apoC-IV, estimated by radioimmunoassay, was in the same range observed in control rabbits injected with physiological saline (data not shown).

The affinity of apoC-IV for lipids was confirmed by its ability to convert small unilamellar vesicles of dimyristoylphosphatidylcholine to discoidal particles. These discs were, however, considerably larger than those produced with rabbit apoE (Fig. 4).

Molecular Cloning of ApoC-IV—Despite its prominence in normal rabbit VLDL, no component resembling apoC-IV was evident upon Coomassie Blue staining of isoelectric focusing electrophoreograms of apoVLDL from rats, mice, or humans and in these species no immunoreactive protein was found in lipoprotein fractions by radioimmunoassay with apoC-IV anti-

sequence (data not shown). To learn more about the origin of apoC-IV, we used an RT-PCR-based cloning strategy to isolate a cDNA encoding apoC-IV. This strategy would encompass cDNAs in rabbit liver of less than 500 bp. To increase the likelihood of obtaining cDNAs from rare transcripts, we used asymmetric PCR in the first round amplification, which gave only a faint smear on ethidium bromide-stained agarose gels.

In the subsequent amplifications with a nested primer, we obtained a band of ~360 bp. Sequencing this band revealed a typical 3'-region of a cDNA containing a poly(A) tail, a polyadenylation signal and a stop codon, together with an open reading frame encoding 90 amino acids beginning with residue 8 of the N-terminal sequence (Fig. 5). The amino acid composition of the predicted polypeptide was in excellent agreement with that determined by amino acid analysis of apoC-IV (Table III). The 5'-cDNA sequence (277 bp) was subsequently obtained by 5'-RACE cloning, as described under "Materials and Methods," which displayed a typical 5' portion of a full-length cDNA containing predicted residues 1–7 of the N-terminal amino acid sequence of apoC-IV (Fig. 5). With this information, the full-length cDNA of apoC-IV was then cloned and sequenced. The sequence was identical except for T rather than C at positions 234 and 398 observed in the 3'-cloning. The earlier result probably reflects an error of incorporation during PCR amplification.

Sequence Analysis of Full-length ApoC-IV cDNA—ApoC-IV cDNA comprises 555 nucleotides, and the decoded cDNA sequence revealed an open reading frame of 372 nucleotides encoding a protein of 124 residues with a molecular mass of 14,097 Da (Fig. 5). The sequence encodes a typical signal pep-
Apolipoprotein C-IV, a New Component of Rabbit Lipoproteins

The open reading frame of 124 amino acids, in one-letter code for the predicted amino acids, is indicated below the nucleotide sequence. A translation initiation codon, ATG, is indicated by the predicted amino acids, is indicated below the nucleotide sequence. A comparison of amino acid residues in apoC-IV, as determined by amino acid analysis with that predicted from cDNA sequencing, is shown in Table III. The values for amino acid composition of apoC-IV are means from 8 analyses of various isoforms separated by preparative isoelectric focusing electrophoresis and ion exchange chromatography. It was assumed that apoC-IV contains 97 amino acids including 5 tryptophan residues, which were not analyzed by amino acid analysis. The value for half-cystine is based upon a single analysis.

| Amino acid | Determined | Predicted |
|------------|------------|-----------|
| Lysine     | 2.8        | 2         |
| Histidine  | 1.0        | 1         |
| Arginine   | 8.9        | 9         |
| Aspartic acid + asparagine | 8.2 | 8 |
| Threonine  | 6.4        | 7         |
| Serine     | 6.3        | 6         |
| Glutamic acid + glutamine | 12.0 | 12 |
| Proline    | 9.3        | 10        |
| Glycine    | 8.3        | 8         |
| Alanine    | 6.0        | 5         |
| Cysteine   | 1.9        | 2         |
| Valine     | 4.1        | 4         |
| Methionine | ND*        | 0         |
| Isoleucine | 0.5        | 0         |
| Leucine    | 13.1       | 14        |
| Tyrosine   | 2.2        | 2         |
| Phenylalanine | 2.2  | 2          |
| Tryptophan | NA         | 5         |

* ND, not detected; NA, not analyzed.

tide of 27 residues and a mature protein of 97 residues (as indicated by the N-terminal amino acid sequence) with a molecular mass of 11,020 Da. The sequence context about the first AUG, GAA AUG is among the most commonly found for functional initiation codons in vertebrates (43). The apoC-IV cDNA sequence contains 97 nucleotides of the 5'-untranslated region composed of nine GGGACAG(A/G) repeats, and 65 nucleotides of the 3'-untranslated region with the polyadenylation signal, AATAAA, 21 nucleotides upstream from beginning of the poly(A) tail.

The amino acids at positions −3 and −1 upstream of the N terminus of the mature protein (Val and Cys, respectively) fit well into the “(−3, −1)” rule for the signal cleavage site of eukaryotes (44). The calculated pl of the mature protein is 5.05. Amphipathic α-helix analysis (45–47) of the predicted amino acid sequence indicates two possible regions (residues 48–71 and 92–113) of the protein that may be responsible for lipid binding (Fig. 6). The sequence lacks the N-glycosylation consensus site NX(S/T), but there are several candidate Ser and Thr residues for O-glycosylation. Two cysteines are present near the C terminus.

DataBank (Swiss-prot, EMBL, and GenBank) searching revealed significant homology between apoC-IV cDNA (146–307 bp) and the 5'-flanking regions of cynomolgus monkey apoC-II (86–247 bp, 77% (48)) and mouse apoC-II (114–244 bp, 82% (49)). This 5'-flanking region of the mouse apoC2 gene has been sequenced recently (9) and contains a novel gene (the “ACL” gene) within the mouse apoE-apoC1-apoC2 gene cluster on chromosome 7. The complete nucleotide sequences of rabbit apoC-IV and mouse ACL cDNA shared 64% homology and the predicted amino acid sequences shared 62% amino acid identity.

Northern Analysis—Northern hybridizations were performed with total RNA isolated from several rabbit tissues. Only a single band of about 0.6 kb of the transcript from liver was detected (Fig. 7), suggesting that our cloned cDNA is complete, and that there is only one major form of mRNA in liver. Our rabbit cDNA failed to recognize any transcript in RNA from human adrenal or HepG2 cells (not shown).

Southern Analysis of Rabbit Genomic DNA—Fig. 8 shows the pattern of hybridization obtained when 537 bp of apoC-IV cDNA was used to probe rabbit genomic DNA digested with four restriction enzymes. In all cases, a single band was found, suggesting a simple gene structure, consistent with the presence of a single copy of the apoC-IV gene per haploid rabbit genome.

Expression of ApoC-IV—E. coli cells, transformed with the expression vector containing the apoC-IV cDNA, but not the vector alone, synthesized proteins of apparent molecular mass of 16 kDa and 14 kDa for the 124-amino acid and 97-amino acid constructs, respectively, which were recognized by guinea pig anti-C-IV (Fig. 9A, lanes 4 and 6). Evidently, the mobility of both forms of apoC-IV is anomalous, like that of apoC-IV isolated from VLDL (Fig. 9A, lanes 2 and 7). To explore this further, we expressed full-length apoC-IV cDNA in mammalian cells.

In cell extracts from transfected COS-1, COS-7, Chinese hamster ovary (CHO), and HepG2 cells, we observed several nonspecific bands around 10–18 kDa in Western blot analysis with our monoclonal antisera (data not shown). Since apoC-IV was isolated from blood plasma and should be detected in medium of transfected cells as a secreted protein, we used the medium from cells to assess the expression of apoC-IV cDNA. Fig. 9B shows that the immunoisolated proteins from medium of CHO cells transfected with pMT2 vector harboring apoC-IV cDNA yielded a specific band of about 16 kDa (lane 4). After desialylation, this band shifted to about 14 kDa (lane 3). Thus, the mobility of asialoglycosylated apoC-IV was similar to that of the unglycosylated protein. Furthermore, the apparent molecular mass of the expressed apoC-IV cDNA product before and after desialylation was similar to the sialylated and desialylated forms of apoC-IV from rabbit apoVLDL, respectively, and after desialylation was similar to the sialylated and desialylated forms of apoC-IV from rabbit apoVLDL (Fig. 9C, lanes 3 and 7).

DISCUSSION

We have identified a novel protein associated with lipoproteins in blood plasma of rabbits. Other proteins, such as serum

Fig. 5. Nucleotide and deduced protein sequences of the cDNA for apoC-IV. The cDNA was isolated from rabbit liver by PCR screening. The open reading frame of 124 amino acids, in one-letter code for the predicted amino acids, is indicated below the nucleotide sequence. A translation initiation codon, ATG, is indicated by bold letters. Positions of oligonucleotides used for PCR amplifications and their orientations with respect to the sense strand are indicated by arrows. * represents the stop codon. The polyadenylation signal is underlined. +1 indicates the start of the mature protein determined by protein sequencing.

Table III

Comparison of amino acid residues in apoC-IV, as determined by amino acid analysis with that predicted from cDNA.

In cell extracts from transfected COS-1, COS-7, Chinese hamster ovary (CHO), and HepG2 cells, we observed several nonspecific bands around 10–18 kDa in Western blot analysis with our monoclonal antisera (data not shown). Since apoC-IV was isolated from blood plasma and should be detected in medium of transfected cells as a secreted protein, we used the medium from cells to assess the expression of apoC-IV cDNA. Fig. 9B shows that the immunoisolated proteins from medium of CHO cells transfected with pMT2 vector harboring apoC-IV cDNA yielded a specific band of about 16 kDa (lane 4). After desialylation, this band shifted to about 14 kDa (lane 3). Thus, the mobility of asialoglycosylated apoC-IV was similar to that of the unglycosylated protein. Furthermore, the apparent molecular mass of the expressed apoC-IV cDNA product before and after desialylation was similar to the sialylated and desialylated forms of apoC-IV from rabbit apoVLDL, respectively, and after desialylation was similar to the sialylated and desialylated forms of apoC-IV from rabbit apoVLDL (Fig. 9C, lanes 3 and 7).

DISCUSSION

We have identified a novel protein associated with lipoproteins in blood plasma of rabbits. Other proteins, such as serum
amyloid A, which is secreted in large amounts as part of the acute phase response (50), associate with lipoproteins and many others, such as complement 4B-binding protein (51, 52), have been found to associate with lipoproteins to a limited extent. The novel protein, which we have designated as apoC-IV, appears to be a bonafide apolipoprotein by several criteria: 1) it is present primarily or exclusively in association with lipoproteins; 2) it binds to phospholipids, converting unilamellar liposomes of dimyristoylphosphatidylcholine to discoidal particles; 3) it is present in measurable quantities at all times and does not appear to be a typical acute phase reactant; 4) it contains two regions predicted to form amphipathic helices of the type found in most soluble apolipoproteins. Based on its molecular weight and distribution mainly in VLDL and HDL, apoC-IV appears to fall into the group of C apolipoproteins. Like apoC-III, it is a sialoglycoprotein, present in several isoforms.

The content of apoC-IV among rabbit C apolipoproteins is appreciable, as judged from Coomassie Blue staining and ion exchange chromatography, the mass of apoC-IV appears to be substantially greater than that of apoC-II and similar to that of apoC-I and apoC-III.

The N-terminal sequence of apoC-IV, with 5 prolines among the first 12 amino acids, gave no clue to its provenance. However, its structure otherwise is typical with respect to its signal peptide and the presence of amphipathic helical sequences downstream. As with other C apolipoproteins, apoC-IV mRNA is expressed mainly (possibly exclusively) in the liver. As predicted, it is evidently secreted from transfected mamalian cells as a sialoglycoprotein. The apparent molecular mass of sialylated apoC-IV is about 15,500 Da; that of asialoC-IV (approximately 14,000 Da) in SDS-PAGE is about 3,000 Da higher than that of the aminoacyl chain (11,020 Da). This anomalous behavior is evidently not related to the presence of one or more O-linked carbohydrate moieties (Fig. 9). The proline-rich N-terminal region may contribute to the reduced mobility.

The protein sequence of apoC-IV showed little homology with that of other apolipoproteins, but, surprisingly, the cDNA sequence showed a striking homology with the 5'-flanking region of monkey and mouse apoC-II (48, 49). The basis for this homology has become clear from the recent work of van Eck, Hoffer and associates (9), who initially found two additional exons 5' to the start site of the mouse apoC2 gene (49). Evidence for a similar gene (the "ECL" gene) was found earlier in the rat by Shen and Howlett (53), in the region of the apoE gene, but with an orientation apparently opposite to that found in the mouse. Van Eck and associates (9) have recently shown that these exons are contained within a novel gene in the apoE,
apoC1, apoC2 gene cluster on mouse chromosome 7, the orientation of which is identical with that of the other three genes in this cluster (Fig. 10). This gene, which they designated the apoC2-linked (ACL) gene, is composed of three exons and is expressed in liver as a transcript of 473 bp. It encodes a putative protein of the precise length of apoC-IV (124 amino acids). The mRNA for the putative mouse protein, like apoC-IV, is expressed solely in the liver. From the sequence of the C-terminal portion of the putative signal peptide (9), we predict that the site of signal peptide cleavage would be the same as for apoC-IV (following residue 27). The overall homology of amino acids 1–27 is 68%. Except for the first 7 amino acids (Fig. 5), the predicted sequence of the mature mouse protein also shows striking homology to apoC-IV, 71% for amino acids 35–124 and when conservative substitutions are included, 80% (Fig. 10). Essentially the same amphipathic helical regions are predicted for the putative mouse protein encoded by the ACL gene and apoC-IV (Fig. 6).
the polar face, for which only the predicted human sequence contains a positively charged amino acid (Lys-110) (Fig. 6). A 28-mer amphipathic helix (residues 47–74) differs from the corresponding rabbit and mouse sequences by the presence of a codon for threonine (ACA) at residue 60 (Fig. 10). In all other species (monkey at predicted residue 60 and mouse, rat and rabbit at predicted residue 57), the equivalent codon specifies proline (CCG/GA) which should produce a kink in the helix (55). If this were not the case, the sequences for the species other than the human would predict a similar amphipathic helix in this region (Fig. 6).

The structure of the mouse and human apoC4 genes as well as that of the rabbit2 is similar, with three exons and two introns (two short exons upstream, the first encoding most of the signal peptide, and a longer terminal exon encoding the C-terminal region). The second and third exons each contain an amphipathic helical region. This genomic organization and the localization of the apoC4 gene within the apoE, apoC1, apoC2 gene cluster provide strong support for inclusion of apoC-IV in the apolipoprotein gene family, which includes not only the four C apoproteins and apoE, but also apoA-I, -A-II, and -A-IV (1). Conservation of the genomic and protein structure of apoC-IV across mammalian evolution from mouse to humans strongly suggests that apoC-IV has an important function. It is thus puzzling that the protein apparently is not expressed and secreted at a more substantial level in species other than the rabbit. A low level of secretion in humans has been predicted by Allan et al. (54), based upon a level of expression of apoC-IV mRNA in human liver only about 1% of that of apoC-II. Clearly, further work is needed to clarify the reasons for such variable expression of apoC-IV (and in particular the basis for an important function at some stage of life).

Acknowledgments—We thank Philippe Duchateau for assistance in verifying the N-terminal sequence of apoC-IV, Clive Pullinger for synthesizing several oligonucleotides, Barney Welsh for providing pBluescript I expression vector, and Robert Hamilton for visualization of lipid-protein complexes. We also thank Clive Pullinger and Qian Jin Hu for critically reading the manuscript and John Kane for advice.

REFERENCES

1. Li, W. H., Tanimura, M., Luo, C. C., Datta, S., and Chan L. (1988) J. Lipid Res. 29, 245–271
2. Chapman, M. J. (1980) J. Lipid Res. 21, 789–853
3. Imaizumi, K., Fainaru, M., and Havel, R. J. (1978) J. Lipid Res. 19, 712–722
4. Havel, R. J., Kane, J. P., and Kashyap, M. D. (1973) J. Clin. Invest. 52, 32–38
5. Brown, W. V., and Bagnsky, M. L. (1972) Biochim. Biophys. Res. Commun. 46, 375–382
6. Havel, R. J., Fielding, C. J., Oliveira, T., Shore, V. G., Fielding, P. E., and Egerrud, T. (1973) Biochemistry 12, 1828–1833
7. Windler, E., and Havel, R. J. (1965) J. Lipid Res. 6, 556–565
8. Havel, R. J., Kita, T., Kotte, L., Kane, J. P., Hamilton, R. L., Goldstein, J. L., and Brown, M. S. (1982) Arteriosclerosis 2, 467–474
9. Van Eck, M. M., Hoffer, M. J. V., Havekes, L. M., Frants, R. R., and Hoffer, M. H. (1994) Genomics 21, 110–115
10. Duff, G. L., and McMillan, G. C. (1949) J. Exp. Med. 89, 611–630
11. Saxena, U., Nagpurkar, A., and Mookerjee, S. (1985) Can. J. Biochem. Cell Biol. 63, 1014–1021

2 L.-H. Zhang, L. Kotite, and R. J. Havel, unpublished data.