Cloning and Expression of Human Apolipoprotein D cDNA*

(Received for publication, May 13, 1986)

Dennis Drayna, Christopher Fielding†, John McLean, Bradford Baer, Graciela Castro‡, Elson Chen, Lisa Comstock, William Henzel, William Kohr, Lucy Rhee, Karen Wion, and Richard Lawn

From Genentech, Inc., South San Francisco, California 94080 and the Cardiovascular Research Institute and the Department of Physiology, University of California, San Francisco, California 94143

The amino acid sequence of human apolipoprotein D, a component of high density lipoprotein, has been obtained from the cloned cDNA sequence. The 169-amino acid protein has no marked similarity to other apolipoprotein sequences, but has a high degree of homology to plasma retinol-binding protein and other members of the \( \alpha \)-globulin protein superfamily. Apolipoprotein D mRNA has been detected in human liver, intestine, pancreas, kidney, placenta, adrenal, spleen, and fetal brain tissue. Tissue culture cells transfected with the cloned cDNA secrete material that reacts with anti-apoD antibodies.

Apolipoprotein D (apoD) is a protein component of high density lipoprotein in human plasma, comprising about 5% of total high density lipoprotein (1). It is a glycoprotein of estimated M, \( \sim 33,000 \) (2). ApoD has attracted considerable interest in studies of plasma cholesterol metabolism because of its physical association in plasma with the enzyme lecithin:cholesterol acyltransferase. In different studies using immunoaffinity chromatography, most or all of lecithin:cholesterol acyltransferase was found complexed with apoD (1, 3), suggesting a role for apoD in the activity of the plasma enzyme in lipid binding or transfer. ApoD, in association with lecithin:cholesterol acyltransferase and other factors, may form a complex which is active in the efflux of cholesterol from peripheral tissues and the transport of cholesteryl esters to the liver for catabolism.

**Materials and Methods**

Isolation of Human Plasma ApoD—ApoD was isolated by ultracentrifugal flotation within the density limits 1.12–1.25 g/ml from normal human plasma essentially as previously described (4) except that delipidation was with ethanol/diethyl ether (2:1, v/v) instead of chloroform/methanol. Purification was carried out by repeated hydroxylapatite chromatography in 0.001 M potassium phosphate buffer (pH 8.0) containing 2 M urea. Purity of the final nonadsorbed protein fraction was confirmed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Protein Analysis—Amino acid composition, protein cleavage, peptide separation, and sequence analysis were performed as previously described (6). The amino terminus was deblocked for sequencing by digestion of 200 pmol of apoD with 0.1 \( \mu \)g of pyroglytate peptidase (7). Supernatant proteins from Chinese hamster ovary cells transfected with apoD cDNA-containing plasmids (and control cells) were metabolically labeled with \[^{35}S\]methionine for 6 hr, immunoprecipitated with rabbit antiserum against plasma-derived human apoD (1), electrophoresed in a 7–17% polyacrylamide gel, and autoradiographed (8).

cDNA Cloning—Human adult liver cDNA libraries in \( \lambda gt10 \) were prepared and provided by Axel Ullrich and Lisa Coussens of Genentech, Inc. Recombinant phage were screened on nitrocellulose filters with single oligonucleotide probes ranging from 33 to 90 bases (9). DNA from cDNA and genomic clones was isolated and sequenced as described (9). Both strands of DNA were independently sequenced. A plasmid for expression of apoD was constructed using the expression plasmid \( pS42E \), which expresses hepatitis B surface antigen (17). This plasmid consists of an \( SV40 \) origin and early promoter preceding the viral DNA encoding hepatitis B surface antigen and its transcription termination region. In addition, this plasmid contains the \( pML \) prokaryotic origin of replication of \( am\prime \) gene, allowing its selection and propagation in *Escherichia coli*. After the EcoRI site 5' to the SV40 promoter was destroyed, this plasmid was digested with EcoRI and HpaI, which liberated only the hepatitis B surface antigen gene, and in its place was ligated an EcoRI-NruI-HpaI synthetic linker 17 nucleotides in length. This plasmid was digested with EcoRI to open it at this single synthetic EcoRI site, and the 698-bp EcoRI insert of \( cAPOD.8 \) was ligated in to create the intermediate plasmid \( pS8AH.5 \). \( pS8AH.5 \) was linearized by digestion with \( SstII \) and a 1713-bp \( SstII \) fragment containing the mouse dihydrofolate reductase gene (from plasmid \( pgD \) trunc dihydrofolate reductase, Ref. 8) was ligated in to give the final expression plasmid designated \( pSAH.8 \). \( pS8AH.8 \), therefore, consists of the \( SV40 \) origin and early promoter leading into the apoD cDNA which is followed by the hepatitis B surface antigen transcription termination region. This is followed by a second \( SV40 \) origin and early promoter, leading into the dihydrofolate reductase gene. The dihydrofolate reductase gene is followed by the origin and \( am\prime \) gene of \( pML \), giving \( pS8AH.8 \) a total length of 6599 bp.

RNA Analysis—RNA was prepared (10) for northern blot analysis following electrophoresis in 1.5% agarose gels containing 6% formamide. Nitrocellulose filters were probed with \[^{32}P\]-labeled apoD insert from the cDNA clone \( cAPOD.8 \) as described (9).

**Results**

Partial Protein Sequence, Oligonucleotide Probes, and cDNA Cloning—Purified human apoD protein was subjected to limited microsequence analysis to generate oligonucleotide probes for cDNA clone identification. The amino terminus of the plasma-derived protein was blocked, but protein sequence was obtained from several peptides derived from cleavage with trypsin or cyanogen bromide. Single oligonucleotides were synthesized to represent one possible codon choice for each of the four regions of available peptide sequence. These probes ranged in size from 33 to 95 nucleotides. About 2 million phage from an oligo(dT)-primed human adult liver cDNA library (11) in \( \lambda gt10 \) were screened with triplicate sets of nitrocellulose filters hybridized with various combinations of \[^{32}P\]-labeled oligonucleotide probes (9). Two phage were recovered which produced detectable hybridization with each of the probes apoD.3 (66-mer), apoD.4 (45-mer), and apoD.6

*This work was supported by Genentech, Inc. and by National Institutes of Health Specialized Center for Research in Atherosclerosis Grant HL14237 (to C. F.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J02561.

1 The abbreviations used are: apoD, apolipoprotein D; bp, base pair.
(95-mer). (The recovery of only two apoD clones from 2 million phage was below expectations and will be discussed below.) Rescreening of the library filters with these cDNA clones as probes yielded only sibling copies of the same clones. Phage DNA was prepared, and the cDNA inserts were analyzed by sequencing.

cDNA Sequence—The ~700-bp inserts of the cDNA clones cAPOD.6 and cAPOD.8 were subcloned into the EcoRI site clones as probes yielded only sibling copies of the same clones. ApoD mRNA and cDNA clones are diagrammed in Fig. 1 (Appendix), and the DNA sequence is presented in Fig. 2 (Appendix). The initial dideoxy sequence of cAPOD.6 revealed an insert of 708 bp (not including synthetic linker sequences), whereas cAPOD.8 contained 698 bp. The two cDNA clones differed slightly in length and contained three internal nucleotide differences. Neither clone contained a poly(A) tail. Translation of the DNA sequence reveals a methionine codon at sequence positions 62-64, followed by a continuous open reading frame. The initial methionine codon is preceded by CAAG and followed by G, in fair agreement with the consensus sequence surrounding translation start sites in eukaryotic mRNAs (14). It is followed by a peptide sequence rich in hydrophobic residues, consistent with an amino-terminal secretion signal peptide (15). The translated DNA sequence contains the regions of sequenced protein which were used in preparing the oligonucleotide probes. The derived protein sequence agrees with all 22 residues of the sequenced CNBr fragment represented in probe apoD.3, with all 15 residues of the partial tryptic fragment of probe apoD.4, and with all 32 residues of the partial tryptic fragment of probe apoD.6 (which overlaps apoD.3). This sequence also agreed with the protein sequence of 11 residues used for the apoD.2 probe, even though it did not hybridize significantly during the library screen. Other regions of partial peptide sequence or where peptide data were obtained after cDNA cloning were also found to correspond to the cDNA-derived sequence and are underlined in Fig. 2 (appendix).

The availability of the cDNA sequence suggested a strategy for obtaining amino-terminal protein sequence information. The glutamine residue, 21 amino acids from the initiator methionine, was a likely candidate for several reasons. It comes after the first charged residue after the hydrophobic "core" of the presumed signal peptide and immediately follows glycine, a common final residue of a leader peptide (15). Furthermore, blockage of amino termini is frequently caused by cyclization of glutamine residues. Therefore, plasma-derived apoD protein was treated with pyroglutamate peptidase, which did unblock the amino terminus. Nine residues of amino-terminal sequence were then obtained, which did not hybridize significantly during the library screen. These data were consistent with the 3' untranslated region being incomplete. Neither clone possessed poly(A) tails. The total length of cloned cDNA is 810 bp, which, with a poly(A) tail, is consistent with the mRNA length of ~900 bases.

Expression of Recombinant ApoD—The cloned apoD coding sequences were expressed from the SV40 early promoter by inserting the EcoRI cDNA insert of cAPOD.8 into the vector plasmid pSLH (see "Materials and Methods"). The resulting plasmid, designated pSAHD.8, was introduced into Chinese hamster ovary cells and amplified by methotrexate selection.
Tissue Distribution of ApoD mRNA—All of the apolipoproteins that have been characterized are synthesized in the liver, although other sites of synthesis occur for some of these proteins. Poly(A)+ RNA was isolated from a number of human tissues and cell lines, electrophoresed, blotted onto nitrocellulose, and hybridized with radiolabeled cloned apoD cDNA. Liver RNA contained a hybridizing 900-base species (Fig. 3b). Pancreas, adrenal gland, kidney, small intestine, and placenta also contain hybridizing apoD RNA of the same size. The signal in these tissues is more intense than in liver, indicating that apoD mRNA accounts for a higher percentage of message in these tissues. ApoD mRNA was not detected in white blood cells or in monocyte-like U937 and HL60 cell lines. In addition, apoD mRNA was found in spleen and in fetal brain in amounts comparable to that in intestine (data not shown).

ApoD Homologies—The nucleic acid sequence of apoD cDNA and its predicted amino acid sequence were compared to other known gene and protein sequences present in the GenBank™ and Dayhoff sequence data bases. No significant amino acid homology was noted between the apoD sequence and those of the apolipoprotein A, B, C, and E groups. In addition, application of the Chou and Fasman algorithm (21) predicts less than 5% α-helix in apoD, in contrast to the substantial amounts of α-helix present in apolipoprotein groups A, C, and E (22, 23).

Significant homology was found between apoD and a class of related proteins, including human retinol-binding protein, human α1-microglobulin, unagulate β-microglobulin, rodent α2-microglobulin, and tobacco hornworm insecticytin. These proteins make up the α2-microglobulin superfamily, and their homologies to apoD extend throughout the length of the apoD molecule. Detailed comparison of apoD to human retinol-binding protein (28), for example, shows 5 out of 6 identical residues near the amino terminus (extending to 20 out of 25 homologous residues if conservative substitutions are allowed), 10 out of 16 identical residues in a central region (17 out of 18 for conservative substitutions), and 7 out of 9 residues (16 out of 17 for conservative substitutions) in a more carboxyl-terminal region. The overall amino acid homology is approximately 25%.

DISCUSSION

The apoD protein sequence derived from translation of the cloned cDNA sequence substantially agrees with the amino acid composition of the plasma-derived protein, as determined by two other independent groups (2, 19). The number of amino acids per molecule of mature apoD we derived by composition analysis and by cDNA sequencing (parentheses) are as follows: Asx, 22.4 (23); Thr, 10.1 (11); Ser, 7.4 (7); Glx, 20.8 (19); Pro, 12.0 (12); Gly, 9.7 (6); Ala, 11.9 (10); Val, 12.0 (12); Cys, 4.3 (5); Met, 2.8 (3); Ile, 9.7 (11); Leu, 13.5 (15); Tyr, 5.4 (7); Phe, 5.1 (7); Lys, 11.0 (11); His, 1.8 (2); Arg, 3.6 (4); and Trp, not determined (4). This similarity of the predicted and directly measured amino acid composition, plus the agreement of extended sequences of apoD from peptide sequencing with the cDNA sequence (Fig. 2, Appendix), and the immunoreactivity and similar gel mobility of plasma-derived and recombinant DNA-derived proteins (Fig. 3) demonstrate that the sequence presented here corresponds to full-length native apoD. The accuracy of the DNA predicted protein sequence was enhanced by the sequencing of three independent cDNA clones plus a genomic clone. The difference between the predicted protein mass (29 kDa) and gel mobility reflects the extensive glycosylation of apoD, as reported before (19) and as indicated by our digestion of the purified glycoprotein by N-glycanase (data not shown). In addition, asparagine-linked glycosylation at the two predicted sites was indicated by peptide sequencing.

The tissue distribution of apoD mRNA is distinct from that of most other apolipoproteins. Several apolipoproteins are synthesized primarily in the liver and intestine, which are the tissues known to be most active in the secretion of intact lipoproteins (20). Whereas the liver synthesizes and secretes apoA (24) as well as lecithin:cholesterol acyltransferase, the present studies suggest that other sites are more active in apoD synthesis. The functions of this distribution are not known.

Both the amino acid sequence and the mRNA distribution of apoD place it apart from the other apolipoproteins characterized to date. Comparison of the protein and gene sequence of the major plasma apolipoproteins has shown that several of these are related in that they share a repeated, 21-residue sequence whose consensus describes regions of amphipathic helix separated by proline residues (29). In particular, apolipoproteins A-I, A-IV, and E show a strong relationship likely to be derived from origin in a common ancestral gene (30). Analysis of the sequence of apoD shows no apparent homology with these other apolipoproteins. Whereas it is a relatively minor component of total plasma apolipoprotein, apoD is found completely in association with lipids in plasma in lipoproteins also containing apoA-I, the major protein of human high density lipoproteins (1). Even after treatment with chaotropic solutions such as 3 M NaSCN, apoD remains associated with the major lipids found in high density lipoprotein (19). By accepted criteria, therefore, apoD has the properties of a plasma apolipoprotein. However, since it contains little if any predicted amphipathic helix, its association with lipids probably has a different basis than that found in apolipoproteins such as apoA-I.

In fact, the homology of apoD to retinol-binding protein suggests such a different interaction between apoD and hydrophobic moieties. The crystallographic structure of retinol-containing human retinol-binding protein has been elucidated (27), and it reveals that this protein binds retinol in a pocket delineated by β-sheet structure, designated a flattened cone (28). We suggest that apoD may be associated with lipid via an analogous structure. Most, or essentially all, of the enzyme lecithin:cholesterol acyltransferase in plasma is found associated in lipoproteins containing apoD (1–3). The major sequence homology reported here between apoD and plasma retinol-binding protein seems to strengthen the concept that apoD plays a role in lipid transfer related to the enzyme reaction. It seems possible that apoD functions to bind either the substrates of the lecithin:cholesterol acyltransferase reaction (lecithin or free cholesterol) or the product of the reaction (cholesterol ester). Elucidation of the precise function of apoD and its relation to the enzyme activity will be facilitated by the availability of pure recombinant apoD. In addition, expression of the cloned apoD gene should make it possible to perform in vitro mutagenesis and thereby define regions of the molecule which might affect lecithin:cholesterol acyltransferase activity.
**Acknowledgments**—We thank Parakash Jhurani and Peter Ng for oligonucleotide synthesis, Lucy Rhee for sequencing, Keri Tate and Gordon Vehar for immunoprecipitation analysis and Jeanne Arch and Alane Gray for manuscript and illustration preparation.

**REFERENCES**

1. Fielding, P. E., and Fielding, C. J. (1986) Proc. Natl. Acad. Sci. U.S.A. 77, 3327-3330
2. Albers, J. J., Cheung, M. C., Ewens, S. L., and Tollefson, J. H. (1981) Atherosclerosis 39, 395-409
3. Utermann, G., Menzel, H. J., Adler, G., Dieker, P., and Weber, W. (1980) Eur. J. Biochem. 107, 225-241
4. McConathy, W. J., and Alaupovic, P. (1975) FEBS Lett. 37, 178-182
5. Kostner, G. (1974) Scand. J. Clin. Lab. Invest. 33, Suppl. 137, 19-21
6. Aggerwal, B. B., Henzel, W. J., Moffat, B., Kohr, W. J., and Harkins, R. N. (1985) J. Biol. Chem. 260, 2334-2344
7. Zalut, C., Henzel, W. J., and Harris, C. H., Jr. (1980) J. Biochem. Biophys. Methods 3, 11-30
8. Lesky, L. A., Dowbenko, D., Simonsen, C. C., and Berman, P. W. (1984) Biotechnology 2, 527-532
9. McLean, J., Fielding, C., Drayna, D., Dieplinger, H., Beer, B., Kohr, W., Henzel, W., and Lawn, R. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 2355-2359
10. Cathala, G., Savouret, J., Mendez, B., West, B., Karin, M., Martial, J., and Baxter, J. (1983) DNA (N. Y.) 2, 329-335
11. Ulrich, A., Cousens, L., Hayflick, J., Dull, T., Gray, A., Tam, A., Lee, J., Yarden, Y., Liberman, T., Schlessinger, J., Downward, J., Mays, C., Whittle, N., Waterfield, M., and Seeburg, P. (1984) Nature 309, 418-425
12. Messing, J., Crea, R., and Seeburg, P. H. (1981) Nucleic Acids Res. 9, 309-321
13. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467
14. Kozak, M. (1980) Nucleic Acids Res. 8, 127-142
15. Perlman, D., and Halverson, H. O. (1983) J. Mol. Biol. 167, 391-409
16. Proudfoot, N., and Brownlee, G. (1981) Nature 285, 359-362
17. Crowley, C. W., Liu, C. C., and Levinson, A. D. (1983) Mol. Cell. Biol. 3, 44-53
18. Simonsen, C. C., and Levinson, A. D. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 2495-2499
19. McConathy, W. J., and Alaupovic, P. (1976) Biochemistry 15, 515-520
20. Mahley, R. W., Innerarity, T. L., Rall, S. C., Jr., and Weisgraber, K. H. (1984) J. Lipid Res. 25, 1277-1294
21. Chou, P. Y., and Fasman, G. D. (1978) Annu. Rev. Biochem. 47, 251-276
22. Knott, T. J., Rall, S., Innerarity, T., Jacobson, S., Urdea, M., Levy-Wilson, B., Powell, L., Pease, R., Eddy, R., Naical, H., Byers, M., Priestly, L., Robertson, E., Rall, L., Betsholtz, C., Shows, T., Mahley, R., and Scott, J. (1985) Science 230, 37-42
23. Protein Sequence Data Base, Release 10.1, Aug. 12, 1986, National Biomedical Research Foundation, Washington, D.C.
24. Erickson, S. K., and Fielding, P. E. (1986) J. Lipid Res., in press
25. Colantuoni, V., Romano, V., Bensi, G., Santoro, C., Costanzo, F., Raugei, G., and Cortese, R. (1983) Nucleic Acids Res. 11, 7769-7776
26. Deleted in proof
27. Newcomer, M. E., Jones, T. A., Aqvist, J., Sundelin, J., Eriksson, U., Rask, L., and Peterson, P. A. (1984) EMBO J., 3, 1451-1454
28. Leurent, B. C., Nilsson, M. H. L., Bavi, C. O., Jones, T. A., Sundelin, J., and Peterson, P. A. (1985) J. Biol. Chem. 260, 11476-11480
29. Elshourbagy, N. A., Walker, D. W., Boguski, M. S., Gordon, J. I., and Taylor, J. M. (1986) J. Biol. Chem. 261, 1998-2002
30. Tsao, Y-K., Wei, C-F., Robberson, D. L., Gotto, A. M., Jr., and Chan, L. (1985) J. Biol. Chem. 260, 15222-15231

**APPENDIX**

**FIG. 1. Human apolipoprotein D mRNA.** Human apoD 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; shown as dotted line), leader peptide coding region (hatched box), mature protein coding region (open box), and 3'-untranslated region ending in polyadenylation site (line with bar). Restriction enzyme sites for MspI (M), TagI (T), BamHI (B), and HincII (H) are shown. There are no sites for EcoRI or HindIII in the cDNA. Below this line, the filled horizontal bars indicate the extent of cDNA clones, and the hatched bar depicts the region of a genomic clone (extending beyond the poly(A) site) that was sequenced to elucidate the 3'-untranslated region of the gene. Below this, arrows indicate the extent and direction of cDNA sequencing reactions. Clones cAPOD.6 and cAPOD.8 and the 450-bp genomic fragment were sequenced from each end, whereas cAPOD.16 was sequenced on one strand only, utilizing two oligonucleotide primers.
Cloning and Expression of Human Apolipoprotein D

FIG. 2. Sequence of human apolipoprotein D cDNA. Nucleotides (left of each row) are numbered from the 5' terminus of cDNA clone cAPOD.6. The complete amino acid sequence of apoD is shown above the DNA sequence. Negative amino acid numbers (above residues) refer to the leader peptide, whereas positive numbers refer to the mature protein. The polyadenylation signal hexanucleotide is double-underlined. Predicted N-linked glycosylation sites are overlined. Fragments from which direct amino acid sequence data were obtained are underlined. Circled numbers below these lines indicate sequences used to construct oligonucleotide probes which are discussed in the text. Complete DNA sequence was obtained on the three cDNA clones and the 3' genomic fragment shown in Fig. 1. Where they overlap, the four sequences in complete agreement with the following exceptions: clone cAPOD.8 contains C at nucleotide 8 (shown below line), whereas cAPOD.6 and cAPOD.16 contain T at this position; clone cAPOD.16 begins at nucleotide 5; clones cAPOD.8, cAPOD.16, and the single genomic clone sequenced contain C at nucleotide 446, whereas cAPOD.6 contains T at this position, resulting in a Leu to Phe amino acid substitution at residue 110. (The C to T, Phe codon substitution at this position appears to be a rare variant; two other cDNA clones, a genomic clone, and 18 chromosomes tested by oligonucleotide probing of genomic digests all contain a CTC Leu codon at residue 110.) The 3'-untranslated region corresponds to the genomic sequence and agrees with the cDNA clones with the following exceptions in the C-rich region which lies about 70 nucleotides 3' of the stop codon: in cAPOD.6, the sequence beginning with the T marked with an asterisk is TACCATAAGAC-end of clone (as indicated by the dashed three-base deletion); in cAPOD.8, the sequence is TACACCCACCCC-end of clone; and in cAPOD.16 the sequence contains a C to G substitution (indicated by parentheses). These slight differences may represent polymorphisms or may be the result of artifacts in cDNA cloning occurring in this unusual poly(C) stretch.