ASBMB JOURNAL OF LIPID RESEARCH

review

The apolipoprotein multigene family: biosynthesis, structure, structure–function relationships, and evolution

Wen-Hsiung Li,* Masako Tanimura,* Chi-Cheng Luo,* Santanu Datta,† and Lawrence Chan†,**

Center for Demographic and Population Genetics,* University of Texas, Houston, TX 77030, and Departments of Cell Biology† and Medicine,** Baylor College of Medicine, Houston, TX 77030

TABLE OF CONTENTS

I.	Introduction	245
II.	Historical Aspects of Apolipoprotein Gene Expression and Structure	246
III.	Apolipoprotein mRNA Translation and Post-Translational Modification	246
	Co-translational processing of preapolipoproteins	246
	Post-translational processing of mammalian apolipoproteins	247
	a. ApoA-I	247
	b. ApoA-II	247
	Post-translational modification of apolipoproteins	248
	ApoB-48 synthesis in the small intestine	248
IV.	The Apolipoprotein Genes	249
	Chromosomal localization and linkage analysis	249
	a. ApoA-I/C-III/A-IV	249
	b. ApoE/C-I and apoC-II	249
	c. ApoA-II	249
	d. ApoB-100	249
	e. ApoD	249
	Structural organization of the apolipoprotein genes	249
	Structure of the human apolipoprotein B gene	251
	Structure of the human apolipoprotein D gene	251
V.	Structural Features of Apolipoproteins	252
	Lipid binding domains of apolipoproteins	252
	Internal repeats in apoA-I, A-II, A-IV, C-I, C-II, C-III, E, and LAL1	253
	ApoB-100 structure and internal repeats	256
	ApoD structure: lack of internal repeats	257
VI.	Evolution	257
	Rates of nucleotide substitution	258
	Evolutionary relationships among apoA-I, A-II, A-IV, C-I, C-II, C-III, E, LAL1, and LAL2	260
	Are apoB and apoD related to other apolipoproteins?	261
	Similarity of apolipoproteins to subdomains of LCAT	262
VII.	Relative Rates of Evolution and Functional Aspects of Apolipoprotein Structure	262
	ApoC-II	263
	ApoC-III	263
	ApoA-II	263
	ApoA-I	263
	ApoA-IV	263
	ApoE	263
	ApoB-100	264
VIII.	Concluding Remarks	264
IX.	References	265

The plasma lipoproteins are water-soluble macromolecular complexes that transport various lipids in the blood and other tissue fluids. Traditionally, they have been separated by ultracentrifugal flotation into four major density classes: chylomicrons, very low density lipoproteins (VLDL), low density lipoproteins (LDL), and high density lipoproteins (HDL). The incidence of coronary heart disease has been positively correlated with high levels of LDL cholesterol and negatively correlated with high levels of HDL cholesterol.

The protein components of plasma lipoproteins are known as apolipoproteins. The major function of apolipoproteins is lipid transport in the intravascular and extravascular compartments. Many apolipoproteins have, in addition, acquired highly specialized functions. For example, Goldstein and Brown (1) have shown that LDL interacts with cell-surface receptors, thereby regulating cholesterol biosynthesis and metabolism. The major apolipoprotein in LDL is apolipoprotein B (apoB). Recently, apoB was found to be heterogeneous in size and function. Normal VLDL and LDL contain a species of the protein with an apparent

Abbreviations: VLDL, very low density lipoprotein; LDL, low density lipoprotein; HDL, high density lipoprotein; LPL, lipoprotein lipase; LCAT, lecithin:cholesterol acyltransferase; LAL1 and LAL2, lamprey apolipoproteins.

Downloaded from www.jlr.org by guest, on June 19, 2012

M_r of 549,000, designated apoB-100; whereas thoracic duct lymph chylomicrons contain another distinct protein, B-48, with an apparent M_r of 264,000 (2, 3). ApoB-48 is synthesized in the intestine (4) and is important in chylomicron metabolism, while apoB-100 is synthesized in the liver (4) and is an important determinant in the binding of LDL to the LDL receptor (5). ApoE also confers receptor binding capacity to lipoprotein particles to both the LDL receptor (6) and a specific apoE receptor (7, 8). ApoC-II activates lipoprotein lipase (LPL) and is important in chylomicron and VLDL metabolism (9, 10). Conversely, apoC-III seems to inhibit the apoC-II activation of LPL (11, 12). It also prevents the uptake of apoE-containing lipoproteins by liver cells (13, 14). ApoA-I (15), and possibly apoC-I (16) and apoA-IV (17), are thought to activate lecithin:cholesterol acyltransferase (LCAT). Most of the cholesteryl esters in human LDL and HDL appear to be formed by the action of LCAT (18).

While the structure and metabolic fate of the plasma lipoproteins have been the subject of intense investigation for over 50 years, it is only in the last 20 years that the primary structure of apolipoproteins has been elucidated, and in the last 10 years that the biosynthesis of these proteins has been studied at the molecular level, first in a model system in an avian species (19–21), then in mammalian species, including man (4, 22–25). Furthermore, it is only in recent years that we have gained, through statistical analyses of DNA and protein sequence data, substantial understanding of the evolutionary relationships among the apolipoproteins (26, 27). We have also uncovered some of the unique structural features (e.g., internal repeats, 27–31) as well as the stringency of structural requirements in the individual functional domains of the various apolipoproteins (27). The aim of this review is to summarize our current knowledge of the biosynthesis, structure, structure–function relationships, and evolution of the apolipoprotein genes and to suggest further studies that are needed for a better understanding of the apolipoprotein multigene family.

HISTORICAL ASPECTS OF APOLIPOPROTEIN GENE EXPRESSION AND STRUCTURE

In the description of the plasma lipoprotein system, little attempt was initially made to examine the biosynthesis of the apolipoproteins at the molecular level. Studies in the rat indicated that the liver and intestine are the major source of plasma apolipoproteins (32). Some early studies using radiolabeled amino acids as well as quantitation of the net production of lipoproteins indicated that the apolipoproteins in VLDL, LDL, and HDL are synthesized in the liver (33–35), and the intestine (36–38). The addition of actinomycin D, an RNA synthesis inhibitor, diminished the production of lipoproteins, suggesting that RNA synthesis is important for the continued production of the latter (39). A detailed analysis of apolipoprotein mRNA translation was first performed in a model system in the chicken (19). The laying hen develops severe hypertriacylglycerolemia with a massive increase in plasma VLDL. The immature chick and the cockerel or rooster can be induced to produce high levels of VLDL by the administration of estrogen (19–21, 40–44). Two major apolipoproteins make up over 90% of the protein content of plasma VLDL in these animals. They are apoB and apoVLDL-II (19, 45, 46). The latter protein was studied extensively and provided the first model system for the molecular biology of apolipoprotein biosynthesis (for a review, see ref. 21).

The initial studies stimulated research in mammalian systems including humans. As will be discussed later, many of the unique features in apolipoprotein gene structure and expression in the avian system also apply to the mammalian apolipoproteins. Indeed, there is little doubt that apoVLDL-II is evolutionarily related to the apolipoproteins so well characterized in man and other mammals.

APOLIPOPROTEIN mRNA TRANSLATION AND POST-TRANSLATIONAL MODIFICATION

Co-translational processing of preapolipoproteins

Taking advantage of the abundance of apoVLDL-II mRNA in the liver of the estrogen-treated cockerel, Chan et al. (19, 20, 47) studied the translation of the mRNA in a heterologous cell-free system. The radiolabeled immunoreactive apoVLDL-II protein was analyzed on an SDS-polyacrylamide gel. The newly synthesized protein was found to migrate on the gel with an apparent size larger than plasma apoVLDL-II by approximately 2–3 kiloDaltons. By using different radiolabeled amino acid precursors, the purified immunoprecipitated product was sequenced by automated Edman degradation. This analysis showed that the in vitro product, designated preapoVLDL-II, contains a 24 amino acid NH_2-terminal extension of the mature plasma protein. The properties of this extra sequence are very similar to those previously described for signal peptides found in secretory protein precursors (48–51). The hydrophobic residues are predominantly restricted to the interior segment of the peptide, while the charged and hydroxyl-containing residues are located near the ends. The sequence represents the first documentation of a signal peptide in a plasma apolipoprotein precursor. The translocation of preapoVLDL-II represents a unique case of vectorial migration of a protein through the membrane of the rough endoplasmic reticulum (52). ApoVLDL-II is unique when compared to other secretory proteins in that, being an apolipoprotein, it has a structural component that would normally spontaneously bind to a phospholipid matrix such as the rough endoplasmic reticulum outer surface (see the

Downloaded from www.jlr.org by guest, on June 19, 2012

section on Lipid binding domains of apolipoproteins, below). Translocation of this phospholipid-binding protein requires some mechanism to prevent interaction of the amphipathic region with the outer membrane. Whether control is afforded by a specific topogenic sequence (53) or whether the signal sequence regulates the folding of the nascent polypeptide chain (54) and prevents this interaction is unknown at present.

Subsequent studies indicate the existence of the signal peptide in all apolipoprotein precursors: apoA-I (20, 22, 23, 29, 55, 56), apoA-II (24, 57, 58), apo-IV (25), apoE (59, 60), apoC-I (61, 62), apoC-II (62), apoC-III (63), apoB-100 (64, 65), and apoD (66). The sequences of the signal peptides for the human apolipoproteins are presented in **Table 1.** The lengths of the peptides vary between 18 and 27 amino acids, well within the usual limits for eukaryotic signal peptides (50, 51). These peptides encompass a positively charged NH_2-terminal region, a central hydrophobic region, and a more polar C-terminal region that seems to define the cleavage site. Thus, they show characteristics present in other eukaryotic signal peptides.

In addition to the structural similarities of the apolipoprotein signal sequences to other eukaryotic signal sequences, there are additional similarities in structure among the individual apolipoprotein sequences, suggesting that these proteins may be related in evolution (see below).

Post-translational processing of mammalian apolipoproteins

a. *ApoA-I*

When either human or rat apoA-I mRNA is translated in vitro in the presence of signal peptidase, an 18 amino acid signal peptide is cotranslationally removed by the enzyme. Examination of the cleaved product reveals that it has an additional 6 amino acid residue segment attached to the NH_2-terminal of mature plasma apoA-I. This larger protein has been designated proapoA-I (23). The primary structure of the pro-segment has been deduced by both in vitro translation and radiosequencing (23, 67) as well as from the nucleotide sequence of apoA-I cDNA clones (55, 68). In both the human and the rat (22, 67, 69), the pro-segment contains a COOH-terminal Gln-Gln dipeptide, unusual amino acids for protein precursors that are processed proteolytically.

Pulse labeling experiments in the human hepatoma cell line, HepG2, indicate that the hexapeptide is present in greater than 95% of newly secreted apoA-I (23, 70). Observations using rat hepatocytes and enterocytes in culture, as well as perfused liver and intestinal segments indicate that most of the proapoA-I is secreted intact without significant proteolytical processing (22, 71). Thus, in these experimental systems, the proteolytic processing of proapoA-I appears to be an exclusively extracellular event. The presence of proapoA-I in the plasma partly accounts for the heterogeneity of plasma apoA-I isoforms (23, 71). In contrast, Banergee et al. (72) found that, in cultured chick hepatocytes and in HepG2 cells, the proportion of secreted proapoA-I versus mature apoA-I showed wide variations, depending on the hormonal environment and the presence or absence of fetal bovine serum in the medium. In chicken plasma, Banerjee, Mukherjee, and Redman (73) detected only mature apoA-I, but no proapoA-I. However, Yang et al. (74) found that proapoA-I was present in hen plasma at approximately 10% of the concentration of the mature protein. The chicken apoA-I prosegment consists of a hexapeptide with the sequence ArgSerPheTrpGlnHis (74). Thus, the -GlnGln dipeptide in mammals is replaced by -Gln-His, indicating that the requirement for Gln in the cleavage site is not absolute.

The predominant site of extracellular processing of human proapoA-I processing has not been identified. This enzymatic activity can be found in plasma, and appears to have the highest specific activity in HDL, as compared with total plasma, VLDL, and lymph chylomicrons. A series of inhibitor studies indicates that the enzyme is metal-dependent, and is *not* a serine protease (75).

b. *ApoA-II*

By using analytical techniques similar to those applied to apoA-I, immunoreactive apoA-II synthesized in vitro in the

TABLE 1. Amino acid sequence of signal peptides of the human apolipoproteins

A-IV	M	-	-	-	F	L	K	A	V	V	L	T	L	-	A	L	V	-	-	-	-	-	A	V	A	G	-	-	-	A	R	A
A-I	M	-	-	-	-	-	K	A	A	V	L	T	L	-	A	V	L	-	-	-	-	-	F	L	T	G	-	-	-	S	Q	A
E	M	-	-	-	-	-	K	V	L	W	A	A	L	-	L	V	T	-	-	-	-	-	F	L	A	G	-	-	-	C	Q	A
A-II	M	-	-	-	-	-	K	L	L	A	A	T	V	-	L	L	L	T	I	C	-	-	S	L	E	G	-	-	-	-	-	-
C-III	M	-	-	-	-	-	Q	P	R	V	L	L	V	V	A	L	L	A	L	L	-	-	A	S	A	R	-	-	-	-	-	A
C-II	M	-	-	-	-	-	G	T	R	L	L	P	A	-	L	F	L	V	L	L	V	L	G	F	-	E	-	-	-	V	G	Q
C-I	M	-	-	-	-	-	R	L	F	L	S	L	P	-	V	L	V	V	V	L	S	I	V	L	A	G	P	A	P	A	Q	G
B	M	D	P	P	-	-	R	P	A	L	L	A	L	L	A	L	P	A	L	L	L	L	L	L	A	G	-	-	-	A	R	A
D[a]	M	V	M	L	L	L	L	L	S	A	L	A	G	L	F	G	A	A	E	G												

[a]The signal peptide of apoD has not been aligned with those of the other apolipoproteins because it is probably evolutionarily unrelated to them (see sections on The Apolipoprotein Genes, Structural Features of Apolipoproteins and Evolution).

ASBMB JOURNAL OF LIPID RESEARCH JLR
Downloaded from www.jlr.org by guest, on June 19, 2012

presence of signal peptidase was found to contain a pentapeptide attached to the NH_2-terminus of the mature plasma protein. The sequence of this peptide, AlaLeuValArgArg, is identical in the human and rat proapoA-II (24, 27, 57, 58, 62, 76). It resembles most pro-segments in that it terminates with paired basic amino acid residues.

In the human hepatoma cell line HepG2, proapoA-II processing occurs predominantly extracellularly. The cells appear to secrete an enzyme(s) capable of the cleavage. The cleavage activity is not inhibited by serine protease inhibitors, but is blocked by a number of thiol protease inhibitors. Recently, Gordon et al. (77) have shown that a cathepsin B-like protein, secreted by HepG2 cells, can effect such a cleavage.

A number of apoA-II isoforms have been identified in human plasma, but the individual isoforms have not been fully characterized. Since proapoA-II is rapidly processed by a hepatocyte-derived protein, we speculate that insufficient amounts of the precursor would appear in the circulation to contribute to any of the heterogeneity detected in plasma.

The extracellular processing of the primary translation products of the mammalian apolipoprotein mRNAs has recently been reviewed by Stoffel (78) and by Gordon et al. (79).

Post-translational modification of apolipoproteins

All the major apolipoproteins are polymorphic in plasma. Each exists in two or more isoforms which differ in charge or mass and are detected by two-dimensional gel electrophoresis. The structural basis for some of these apparently heterogeneous forms has been worked out for some of the apolipoproteins and includes amino acid substitution, glycosylation, deamidation, proteolytic cleavage, acylation, and phosphorylation. ApoE is an excellent case of protein charge heterogeneity resulting from allelic differences in amino acid sequence (80, 81). In addition, the following apolipoproteins are known to be glycosylated: apoA-II, B, C-III, and E. ApoC-II appears to be secreted in a glycosylated form and is deglycosylated in the circulation (82). The different isoforms of plasma apoA-I are, in part, the products of deamidation (83). Some of the circulating apolipoproteins undergo further proteolytic cleavage, in many cases possibly as intermediate steps in degradation. It is unclear whether in some cases the proteolytic cleavage products are artifacts of isolation or incubation conditions (84). In the case of apoC-II, a minor isoform designated apoC-$II_{1/2}$ by Fojo et al. (82) was found to be a truncated protein missing the first six N-terminal amino acid residues of the major isoform, apoC-II_o.

Using the human hepatoma cell line HepG2, Hoeg et al. (85) have established that, in addition to proteolytic processing, secreted nascent apoA-I is acylated with palmitate. They also suggest that apoA-II, A-IV, B-100, C-II, and C-III all have some degree of fatty acid acylation. The biological consequence of the acylation is unknown. Hoeg et al. (85) speculate that the increased hydrophobicity conferred upon the protein by the fatty acid might facilitate protein-lipid interaction.

In cultured rat hepatocytes, Davis et al. (86) found that newly secreted apoB-48, but not B-100, was phosphorylated. In radiolabeling experiments when the cells were incubated with ^{32}P-orthophosphate, at least 20% of the ^{32}P associated with apoB-48 was in the form of phosphoserine. In vivo labeling experiments confirmed that phosphorylation of apoB-48 occurs in vivo in the rat. Furthermore, ^{32}P-labeled apoB-48 loses its radioactivity upon incubation in rat serum. The physiological significance of apoB-48 phosphorylation in rat liver is unknown.

ApoB-48 synthesis in the small intestine

In adult humans, apoB-100 is synthesized exclusively in the liver, and apoB-48 in the intestine (87). Glickman, Rogers and Glickman (4) found that the early fetal intestine synthesizes mainly apoB-100, and the capacity to synthesize apoB-48 is acquired during maturation of the fetus. ApoB-48 shares antigenic determinants with the NH_2-terminal half of apoB-100 (88, 89) and has a molecular weight approximately 48% of the latter (2, 3). Studies using monoclonal antibodies (90), Southern blotting (91–93), and genomic cloning and sequencing (94, 95) indicate that apoB-48 and apoB-100 are the products of the same gene. However, experiments involving labeled amino acid precursor incorporation and immunochemical identification indicate that apoB-48 is not a post-translational cleavage product of apoB-100 (87). These puzzling observations can be explained by some recent experiments by Powell et al. (96) and Chen et al. (97). They examined apoB-48 synthesis and structure by five different approaches: DNA-excess hybridization of small intestinal mRNA using apoB-100 cDNA probes, direct sequencing of tryptic peptides of apoB-48 purified from chylous ascites fluid, sequencing of cloned intestinal apoB cDNAs, direct sequencing of intestinal mRNAs, and oligonucleotide hybridization of intestinal cellular DNA. ApoB-48 was shown to be the product of an intestinal mRNA that has an in-frame UAA stop codon resulting from a C → U change in the codon CAA encoding Gln-2153 in apoB-100 mRNA (96, 97). Oligonucleotide hybridization of intestinal cellular DNA indicates that, like leucocyte DNA (95), enterocyte DNA had a C in this location (96). The COOH-terminal Ile-2152 of apoB-48 purified from chylous ascites fluid had apparently been cleaved from the initial translation product, leaving Met-2151 as the new COOH-terminus (97). DNA-excess hybridization and sequencing data indicate that ~85% of the intestinal mRNAs terminate within ~0.1–1.0 kilobases downstream from the in-frame stop codon. The other ~15% of the mRNAs have lengths similar to hepatic apoB-100 mRNA, but they also have the same in-frame stop codon (97). The organ-specific introduction of a stop codon into intestinal apoB-48 mRNA is an unprecedented observation, and the exact mechanism is currently under active investigation.

Downloaded from www.jlr.org by guest, on June 19, 2012

THE APOLIPOPROTEIN GENES

The first cloned cDNA and genomic sequences from the apolipoprotein multigene family were isolated from the avian species pertaining to apoVLDL-II (98–103). In the last 3 years, all the major apolipoprotein cDNAs from humans have been cloned and sequenced, and the genomic structures for eight of them, apoA-I, apoA-II, apoA-IV, apoE, apoC-II, apoC-III, apoB-100, and apoD have been published (52, 54, 55, 65, 66, 68, 94, 95, 104–124). In addition, the cDNAs for a number of rodent, canine, and rabbit apolipoproteins have also been cloned and characterized (27, 30, 59, 125–133). In this section, we shall review the chromosomal localization and structural organization of the human apolipoprotein genes.

Chromosomal localization and linkage analysis

The chromosomal locations of some of the apolipoprotein genes have previously been established by classical family studies and linkage analysis. Recently, most of the apolipoprotein genes have been mapped to specific human chromosomes by two newly developed techniques: *1)* Southern blot analysis of somatic cell hybrids involving human and nonhuman cells, and *2)* in situ nucleic acid hybridization on chromosome spreads. In somatic cell hybrid analysis, hybrid cells formed between human cells and nonhuman (usually rodent) cells are established in vitro. In human–rodent hybrids, there is preferential loss of human chromosomes. It is thus possible to identify human chromosomes that are retained or lost in the hybrids and to correlate them with the presence or absence of a specific apolipoprotein gene detected by Southern blot analysis (134). In situ nucleic acid hybridization was first used to detect repetitive sequences on DNA. Subsequent modifications of the original method allowed the localization of single-copy genes (such as apolipoprotein genes) in mitotic chromosome preparations (135).

The chromosomal localizations of the major apolipoprotein genes are presented in **Table 2,** and are summarized below.

a. *ApoA-I/C-III/A-IV*

The apoA-I, apoC-III, and apoA-IV genes were found to occur in a cluster within a span of about 22 kilobases (63, 106, 136). These genes have been localized to human chromosome 11 by the technique of Southern blot hybridization of DNA isolated from a panel of human–rodent somatic cell hybrids (137–139). The chromosomal assignment was confirmed by analysis of a hybrid containing a single human chromosome, no. 11. Regional mapping was achieved by using deletion mutants that retained different regions of chromosome 11. Such studies localized the apoA-I/C-III/A-IV gene cluster to the region 11q13 → qter on the long arm of chromosome 11 (137).

It is interesting that in the mouse, the apoA-I gene has been assigned to chromosome 9 (127) in a region highly homologous to a region on the long arm of human chromosome 11 where the human apoA-I gene resides.

TABLE 2. Chromosomal localization of the human apolipoprotein genes

ApoA-II	Chromosome 1
ApoB-100	Chromosome 2
ApoD	Chromosome 3
ApoA-I/C-III/A-IV	Chromosome 11
ApoE/C-I, C-II	Chromosome 19

b. *ApoE/C-I and apoC-II*

The human genes for apoE and apoC-I were found to reside within 4 kb of each other (140). By the technique of somatic cell hybridization, these genes (105, 107, 140, 141), as well as that for apoC-II (108, 140, 142), were found to be on human chromosome 19. Pedigree analysis indicates that the apoE and apoC-II gene loci are tightly linked (143, 144).

These studies using molecular biological techniques confirm previous investigations by pedigree analysis that showed that the apoE gene is linked to the complement component *C3* locus on chromosome 19 (145). The gene that is associated with familial hypercholesterolemia (146) or its structural equivalent, the low density receptor (*LDLR*) locus (147), has also been mapped to chromosome 19. The exact location of the apoE/C-I and apoC-II genes on chromosome 19 has not been determined, though it is somewhere close to the centromere.

c. *ApoA-II*

The gene for apoA-II has been mapped to human chromosome 1 in the region 1p21 → 1qter by somatic cell hybrid analysis (57, 141, 148). Interestingly, in the mouse, apoA-II has been mapped to chromosome 1 and is linked to a gene that determines HDL structure (127).

d. *ApoB*

ApoB-100 and apoB-48 are the products of a single gene. The recent availability of apoB-100 cDNA clones has allowed the assignment of this gene to the short arm of human chromosome 2 by the method of somatic cell hybridization (91, 92, 116, 118). In situ nucleic acid hybridization further localized the gene to the region p23–p24 (93, 116).

e. *ApoD*

The apoD gene has been mapped to human chromosome 3 by three independent procedures: *i)* dot-blot hybridization of DNA from flow-sorted human chromosomes; *ii)* Southern blotting to DNA from human–hamster hybrid cell lines; and *iii)* in situ nucleic acid hybridization to metaphase chromosomes (124). By the last procedure, 41% of all the hybridizing ^{3}H-labeled apoD cDNA probes (appearing as silver grains on the autoradiographs) cluster on the distal long arm (3q26.2 → qter).

Structural organization of the apolipoprotein genes

Despite the distribution of the apolipoprotein genes over five chromosomes, the genomic structures for six of the

ASBMB JOURNAL OF LIPID RESEARCH JLR
Downloaded from www.jlr.org by guest, on June 19, 2012

TABLE 3. Structural features of apolipoprotein genes

Genes	Length (Codons) of Coding Region: Signal Peptides	Prosegments	Mature Peptides	Total	Location of Introns: Intron 1 Upstream from AUG Preceding Nucleotide No.	Intron 2 between 1st and 2nd position of Codon No.[a]	Intron 3 between 2nd and 3rd position of Codon No.[a]
Human A-I	18	6	243	267	−20	−4	+49
Rat A-I	18	6	235	259	−19	−4	+48
Human A-II	18	5	77	100	−24	−1	+44
Rat A-II[b]	18	5	79	102	NA[c]	−1	+44
Human A-IV	20	0	377	397	[d]	−4	+39
Rat A-IV	20	0	371	391	[d]	−4	+39
Mouse A-IV	19	0	375	394	[d]	−4	+39
Human C-I[b]	26	0	57	83	NA	NA	NA
Human C-II	22	0	79	101	−13	−4	+50
Dog C-II[b]	22	0	79	101	NA	−4	+50
Human C-III	20	0	79	99	−13	−2	+40
Dog C-III[b]	20	0	80	100	NA	−2	+40
Human E	18	0	299	317	−23	−4	+61
Rat E	18	0	293	311	−23	−4	+53
Mouse E[b]	18	0	293	311	NA	−4	+53
Chick VLDL-II	24	0	83	107	−39	−4	+50

[a]Taking the first codon after the signal peptide cleavage site as codon no. +1. Thus, for the mature peptide regions of apoA-I and apoA-II, the numbers include the 6 and 5 codons of the prosegment respectively.
[b]Genomic sequences are not available for these genes. However, the intron/exon junctions can be inferred from the corresponding cDNA sequences and the human genomic sequences because of the extremely high degreee of homology at these regions.
[c]NA, information not available.
[d]Intron 1 is missing in the apoA-IV genes.

genes studied to date show remarkable similarities. The characteristics of these genes (including some counterparts in the rat and the apoVLDL-II gene in the chicken) are summarized in **Table 3** and **Fig. 1**; the apoB-100 and apoD genes show significant differences and will be considered later. With the exception of apoA-IV, there are four exons and three introns in each gene. The introns appear to have very similar locations: intron-I interrupts the 5' untranslated region of the gene; intron-II interrupts a coding region of the gene very close to the signal peptidase cleavage site;

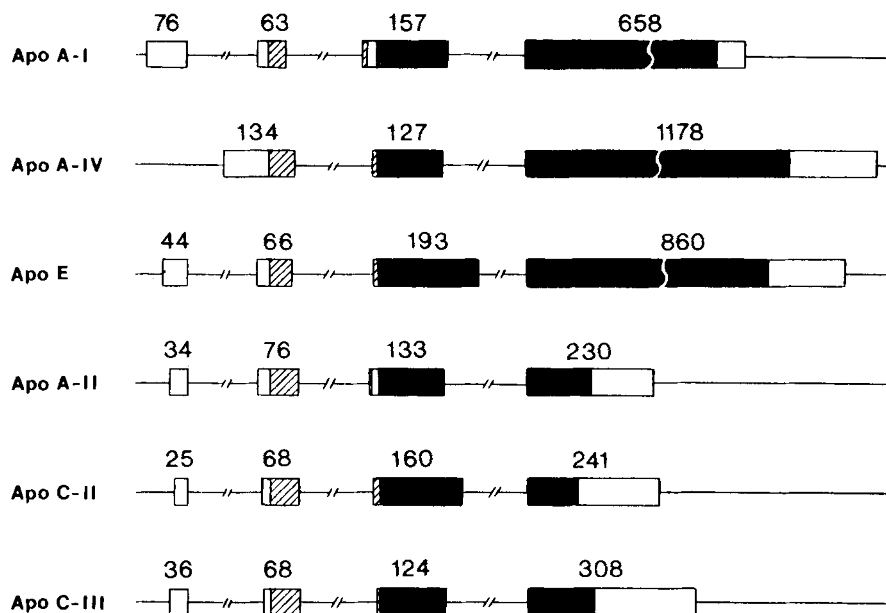

Fig. 1. Structural organization of the human apolipoprotein A-I, A-II, A-IV, C-II, C-III, and E genes. Transcription is from *left* to *right*. The *wide bars* represent the exons, and the *thin line* the 5' flanking region, introns, and 3' flanking region of the respective genes. The *wide bars* are divided into several regions: the *open bars* at the two ends represent the 5' and 3' untranslated regions, the *hatched bars* the signal peptide regions, and the *solid bars* the mature peptide regions of the respective genes. In apoA-I and apoA-II, the prosegment is represented by a *narrow open bar* between the signal peptide and mature peptide region. The *numbers above* the exons indicate the length (number of nucleotides) of the exons. The lengths of the exons are drawn to scale, except for the last exons in apoA-I, E, and A-IV.

Downloaded from www.jlr.org by guest, on June 19, 2012
ASBMB JOURNAL OF LIPID RESEARCH JLR

and intron-III interrupts the part of the gene encoding the mature peptide. ApoA-IV differs from the other genes in that the intron in the 5' untranslated region is missing. However, the other two introns in this gene are present at the same locations as in the other apolipoprotein genes. The lengths of the first three (two for apoA-IV) exons are very similar among genes, and the differences in the total length of mRNAs are accounted for mainly by differences in the length of exon 4 (exon 3 for apoA-IV). This striking similarity in the structure of the six apolipoprotein genes in three different species supports the hypothesis that these genes have arisen from a common ancester, and that the individual apolipoprotein genes have evolved through partial and complete gene duplications (see below).

As in the case of many other eukaryotic genes (e.g., collagen [149], albumin [150], immunoglobulin [151], ovomucoid [152], rhodopsin [153], 3-hydroxy-3-methylglutaryl coenzyme A reductase [154], and pyruvate kinase [155] genes), the introns tend to interrupt the genes at the junctions defined by specific domains as well as by well-formed structures (156). For example, intron-II interrupts the gene close to the signal peptidase cleavage site, either at regions of random strucutre (apoC-II and apoC-III), at β-turn (apoA-I, apoE, and apoVLDL-II), or at the junction between two conformations (i.e., transition between short, probably unstable β-sheet and α-helical region, apoA-II). In these instances, therefore, not only do introns II and III separate the polypeptides into distinct functional domains (156), but they also seem to have preference for the surface of proteins as in the case of many other proteins (157). However, by Chou-Fasman analysis (158), intron-III might interrupt α helical structures in apoA-I and apoE.

Gilbert (156) speculated that the exon-intron structure of eukaryotic genes might be a record of their evolutionary history: these genes evolved by exploiting RNA splicing to recruit and combine segments of coding sequences. Thus, in many instances, the coding portions of these genes are broken into pieces encoding distinct functional domains. Inspection of the apolipoprotein genes clearly indicates that intron-II (intron-I for apoA-IV) marks the demarcation between the signal peptide and the mature protein. To examine the relative position of the last intron with respect to the conformational domains of the mature protein, one can analyze the latter by the mean helical hydrophobic moment $<\mu_H>$ of Eisenberg, Weiss, and Terwilliger (159). This parameter not only gives a measure of helicity, but also takes into consideration the amphiphilicity of a helix. Inspection of **Fig. 2** reveals that this intron consistently interrupts the mature polypeptide at, or very close to, a minimum value of $<\mu_H>$. It should be noted that the *arrows* in this figure indicate the intron position at a site which marks the end of the first domain, e.g., for apoA-II, the *arrow* marks the domain which ends at residue 39 of the mature protein, i.e., at the midpoint of the 11-residue peptide that defines the conformation of residues 29-39, or residue 34. For the longer apolipoproteins, apoA-I, apoE, and apoA-IV analyzed by this method, in addition to the minimum values of $<\mu_H>$ marked by the last intron, other minima seem to appear periodically toward the carboxyl-terminal end of the protein. These subsequent domains are not interrupted by introns and seem to define internally repeated sequences in these proteins (27, 28, 103, 160). The exon-intron organization of the apoC-I gene is not known. However, a clear minimum value of $<\mu_H>$ is identified by the 11-amino acid domian that ends at amino acid number 39 (see Fig. 2, *open arrow*). Furthermore, by sequence alignment, this amino acid residue is at an analogous location where an intron interrupts all the other apolipoprotein genes (27; see Fig. 3a and the section on Internal repeats in A-I, A-II, A-IV, C-I, C-II, C-III, and E, below). We speculate that apoC-I likely has an intron between the second and third nucleotides of the codon specifying amino acid residue number 39.

In addition to sharing a common basic structure with respect to exon-intron organization, the apolipoprotein genes also share significant homology in the coding regions (see below). Moreover, examination of the 5' flanking regions of these genes reveals sequences of high homology that are GC-rich and 14-27 nucleotides in length situated within 500 nucleotides upstream from the respective cap sites (112). These regions might represent structures conserved during evolution and/or regions of functional significance; e.g., they might be regulatory sequences which render the respective genes responsive to changes in cellular lipids, or some lipid-protein complexes, or other common intermediates in apolipoprotein gene expression.

Structure of the human apolipoprotein B gene

Compared to the other apolipoproteins, the human apoB gene shows many remarkable differences (94, 95). Although spanning 43 kilobases, the total length of the apoB gene is actually relatively small for the size of its gene product, which is the largest monomeric protein known. The exons together comprise about one-third of the gene, and the total length of the 28 introns is only about twice that of the exons. Two of the exons (exons 26 and 29) are unusually long: 7572 and 1906 base pairs, respectively. The distribution of introns is extremely asymmetrical, with a concentration in the 5'-terminal one-third of the gene. For example, there is one intron in the 5'-untranslated region, and another 11 introns occur within the first 500 codons, and a total of 19 introns appear within the first 1000 codons.

It is evident that if the apoB gene is evolutionarily related to the other apolipoprotein genes, it has undergone considerable change with time and no longer has the 3-intron/4-exon structure common to most of the other genes.

Structure of the human apolipoprotein D gene

The apoD gene has been recently cloned and characterized by Drayna et al. (124). The organization of this gene

Downloaded from www.jlr.org by guest, on June 19, 2012

ASBMB JOURNAL OF LIPID RESEARCH

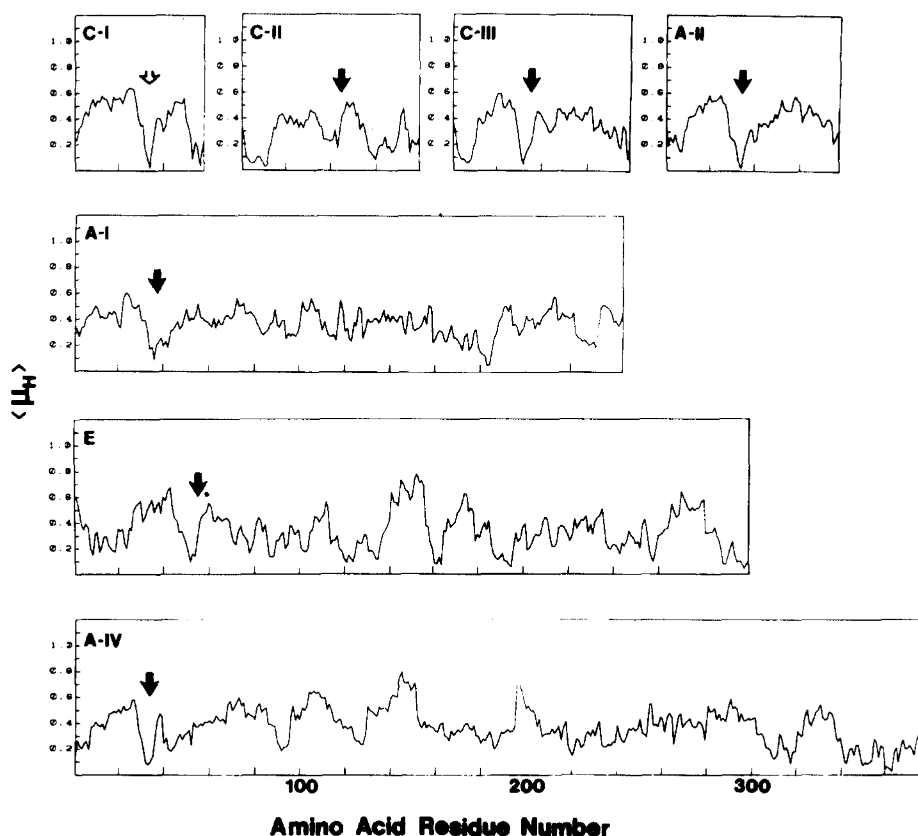

Fig. 2. Predicted secondary structure of human apolipoproteins and intron locations. The helical hydrophobic moment, $<\mu_H>$, was determined by the procedure of Eisenberg et al. (159). Each apolipoprotein was analyzed in 11 residue segments and the $<\mu_H>$ values were plotted at the midpoint of each fragment. The *solid arrows* indicate the positions of the introns. Each arrow is placed at the residue that marks the end of exon 3 (or exon 2 for apoA-IV), which also defines the end of the first domain in this figure. For example, for apoA-II, intron 3 is located at amino acid 39, and the *solid arrow* is placed at the midpoint of the 11-residue peptide which marks the end of the domain encompassed by residues 29-39 or residue 34. The intron location interrupting the mature peptide of apoC-I has not been reported. The *open arrow* marks the position of intron 3 in apoC-I as predicted by sequence alignment. It defines the end of the first domain of the mature peptide region which also coincides with the common 33-codon block defined in Fig. 3 *top* (see below).

is clearly different from those of the other soluble apolipoprotein genes. It has at least one intron in the 5' untranslated region. Since the transcription initiation site has not been defined, additional introns in this region are possible. Unlike the other genes, the signal peptide region in the apoD gene does not contain any intron. Furthermore, in the mature peptide coding region, the apoD gene contains three introns compared to only one in the other soluble apolipoprotein genes. The drastically different structural organization of the apoD gene suggests that it is not a member of the apolipoprotein multigene family (see below).

STRUCTURAL FEATURES OF APOLIPOPROTEINS

Lipid binding domains of apolipoproteins

A basic function of all apolipoproteins is lipid transport. Association of phospholipids with the apolipoproteins A-I and A-II in HDL (161), and apoC-I and C-III in VLDL and HDL results in an increase in α-helical content of the protein as measured by circular dichroism (162). Further, it was shown that some of the tryptophan residues are exposed to solvent in delipidated apoC-I but are buried in a nonpolar environment when the apolipoprotein is associated with phospholipids (163). Based on these observations, and on model-building of apolipoproteins with known sequences, Segrest et al. (164) proposed that these specialized proteins contain amphipathic helical regions that interact with the lipid. An important feature of the amphipathic helix is the presence of two clearly defined faces, one hydrophobic that is inserted between the fatty acyl chains of the phospholipid molecules, and the other hydrophilic that interacts with the phospholipid head groups and the aqueous phase. Another feature of the model is the distribution of the charged amio acid residues on the polar face. The negatively charged residues such as Glu and Asp tend to occur along the center of the polar face, while the positively charged residues, Lys and Arg, are located on the

Downloaded from www.jlr.org by guest, on June 19, 2012
ASBMB JOURNAL OF LIPID RESEARCH

lateral edges of the polar face. Subsequently, similar amphipathic helical regions have been identified in all the major apolipoproteins. Short stretches of similar conformations are also found on the surface of many globular proteins (165).

The lipid-protein interaction between lipids and apolipoproteins differs from that between lipids and membrane-spanning proteins. In the latter case, long segments of exclusively nonpolar amino acid residues facilitate the actual insertion of the polypeptide into the hydrophobic environment. In contrast, the hydrophilic residues on one side of the amphipathic helix are in contact with the surrounding aqueous phase, and keep the apolipoproteins at the surface of the lipoprotein particle. Such a location facilitates transfer between lipoprotein particles and interaction with other molecules such as enzymes (e.g., lipases) and specific cell surface receptors.

ApoB-100 is unique among the apolipoproteins in having a relatively high β-sheet content ($\sim 20\%$), both by predictive analysis of its primary sequence (65, 113, 117, 119, 122) and by circular dichroism and infrared absorption studies of LDL (166, 167). Unlike other apolipoproteins, LDL apoB-100 does not transfer among lipoprotein particles. These observations have led to the speculation that apoB-100 binds lipid by hydrophobic β-sheet regions (168). Recently, Yang et al. (169) isolated thirteen peptides from a mixture of total apoB-100 tryptic peptides by exploiting their phospholipid binding properties. Analysis of all these peptides by circular dichroism and by use of a predictive algorithm reveals no evidence of amphipathic helices. Interestingly, although the predictive algorithm suggests that they display substantial β-sheet potential, no spectroscopic evidence of this structure was found. Therefore, the conformation of these lipophilic tryptic peptide regions on intact low density lipoprotein particles is still unknown.

ApoD also differs from the other apolipoproteins in that the cDNA-deduced amino acid sequence contains little, if any, predicted amphipathic helical structure (66).

Internal repeats in A-I, A-II, A-IV, C-I, C-II, C-III, E and LAL1

About 10 years ago, Barker and Dayhoff (26), Fitch (28), and McLahlan (160) independently observed that apoA-I contains multiple repeats of 22 amino acids (22-mer), each of which is a tandem array of two 11-mers. Recent DNA sequencing work has confirmed the existence of a 22-mer periodicity in apoA-I (29). The repeat unit of 22-mer has been suggested to be a structural element that builds an amphipathic α-helix (170, 171; and see above). Recently, Nakagawa et al. (172) proposed that a 44-mer composed of two 22-mers punctuated in the middle by a helix breaker, Pro, rather than the 22-mer itself (173), is the paradigm of lipid-binding domains of apoA-I. They showed that the 44-mer mimics more closely the surface properties and conformations of apoA-I than does the 22-mer. They speculated that the centrally located Pro residue that breaks the α-helix actually keeps the hydrophobic faces in phase and the resulting concavity of the 44-mer is suited for the adsorption of the peptide to the highly curved surface of human plasma HDL_3 (radius 40–50 Å).

In the last few years, the existence of a 22-mer periodicity has also been found in other apolipoproteins, including apoA-II, A-IV, C-II, C-III, and E, and an 11-mer has been found in apoC-I (27, 30, 31, 105, 106, 109, 125). The internal repeats in human apoA-I, A-II, A-IV, C-I, C-II, C-III, and E are shown in **Fig. 3.**

Using the junction between exon 3 and intron 3 as a reference point, Luo et al. (27) obtained an alignment of the genes for the above proteins for the 33 codons upstream from the junction. In this block of 33 codons, the homologies between genes are, in the majority of cases, about 40% or higher at the nucleotide level, and 15% or higher at the amino acid level (27). The homologies become conspicuous when the alignment is shown at the amino acid level, taking into account similarities in biochemical properties between amino acids (Fig. 3 *top*). A very striking feature is that many columns, e.g., columns 8 and 9, are completely or almost completely occupied by hydrophobic residues (green color). In addition, some columns are predominantly occupied by acidic residues (red), some others by indifferent residues (uncolored), and one by basic residues (blue). When the block is divided into three segments (A, B, and C) of 11 columns, the first column of each segment consists predominantly of hydrophobic residues and that in segments A and B, columns 8 and 9 consist almost completely of hydrophobic residues. Secondary structure analysis of the sequence of the 33 amino acids that make up the common block by the method of Eisenberg et al. (159) indicates that they form a well-defined domain at the N-terminus of each apolipoprotein (see Fig. 2). Luo et al. (27) suggested that the three segments in the common block arose from triplication of an 11-mer (11 amino acids). It is clear from Fig. 3 *top* that the similarity between segments A and B is higher than that between segments B and C. Therefore, it appears that segments B and C were first derived from a duplication of 11 codons and later segment A was derived from a duplication of segment B.

When the junction between intron 3 and exon 4 is used as a reference point, a pattern of internal repeats in exon 4 also becomes evident for each of the genes considered (Fig. 3 *bottom*). The fundamental unit of repeat in exon 4 is not an 11-mer, but a 22-mer that is made up of two 11-mers. The main reason for this is that most 11-mers are more similar to the 11-mer one unit removed than to the adjacent 11-mer (27–29). This is most clearly indicated by the fact that almost all prolines (P) appear in the first column when the alignment is made using 22 amino acids as the basic unit of repeat (Fig. 3 *bottom*). This basic structure of 22-amino acids units connected to one another by a proline residue is highly reminiscent of the 22- and 44-amino acid struc-

Downloaded from www.jlr.org by guest, on June 19, 2012

ASBMB JOURNAL OF LIPID RESEARCH

ASBMB

JOURNAL OF LIPID RESEARCH

		A											B											C											
		1	2	3	4	5	6	7	8	9	10	11	1	2	3	4	5	6	7	8	9	10	11	1	2	3	4	5	6	7	8	9	10	11	
A-I	11	V	K	D	L	A	T	V	Y	V	D	V	L	K	D	S	G	R	D	Y	V	S	Q	F	E	G	S	A	L	G	K	Q	L	N	43
A-IV	7	V	A	T	V	M	W	D	Y	F	S	Q	L	S	N	N	A	K	E	A	V	E	H	L	Q	K	S	E	L	T	Q	Q	L	N	39
E	29	A	L	G	R	F	W	D	Y	L	R	W	V	Q	T	L	S	E	Q	V	Q	E	E	L	L	S	S	Q	V	T	Q	E	L	R	61
A-II	7	V	E	S	L	V	S	Q	Y	F	Q	T	V	T	D	Y	G	K	D	L	M	E	K	V	K	S	P	E	L	Q	A	E	A	K	39
C-III	8	L	L	S	F	M	Q	G	Y	M	K	H	A	T	K	T	A	K	D	A	L	S	S	V	Q	E	S	Q	V	A	Q	Q	A	R	40
C-II	18	V	K	E	S	L	S	S	Y	W	E	S	A	K	T	A	A	Q	N	L	Y	E	K	T	Y	L	P	A	V	D	E	K	L	R	50
C-I	7	A	L	D	K	L	K	E	F	G	N	T	L	E	D	K	A	R	E	L	I	S	R	I	K	Q	S	E	L	S	A	K	M	R	39
LALI	10	F	P	D	A	F	W	E	S	F	K	N	V	S	M	E	F	K	K	M	V	H	G	L	Q	T	S	N	I	G	E	H	A	K	42

		A											B												
A-I	44	L	K	L	L	D	N	W	D	S	V	T	S	T	F	S	K	L	R	E	Q	L	G	65	A-I-4
	66	P	V	T	Q	E	F	W	D	N	L	E	K	E	T	E	G	L	R	Q	E	M	S	87	A-I-5
	88												K	D	L	E	E	V	K	A	K	V	Q	98	A-I-6
	99	P	Y	L	D	D	F	Q	K	K	W	Q	E	E	M	E	L	Y	R	Q	K	V	E	120	A-I-7
	121	P	L	R	A	E	L	Q	E	G	A	R	Q	K	L	H	E	L	Q	E	K	L	S	142	A-I-8
	143	P	L	G	E	E	M	R	D	R	A	R	A	H	V	D	A	L	R	T	H	L	A	164	A-I-9
	165	P	Y	S	D	E	L	R	Q	R	L	A	A	R	L	E	A	L	K	E	N	G	G	186	A-I-10
	187	A	R	L	A	E	Y	H	A	K	A	T	E	H	L	S	T	L	S	E	K	A	K	208	A-I-11
	209	P	A	L	E	D	L	R	Q	G	L	L												219	A-I-12
	220	P	V	L	E	S	F	K	V	S	F	L	S	A	L	E	E	Y	T	K	K	L	N	241	A-I-13
A-IV	40	A	L	F	Q	D	K	L	G	E	V	N	T	Y	A	G	D	L	Q	K	K	L	V	61	A-IV-4
	62	P	F	A	T	E	L	H	E	R	L	A	K	D	S	E	K	L	K	E	E	I	G	83	A-IV-5
	84												K	E	L	E	E	L	R	A	R	L	L	94	A-IV-6
	95	P	H	A	N	E	V	S	Q	K	I	G	D	N	L	R	E	L	Q	Q	R	L	E	116	A-IV-7
	117	P	Y	A	D	Q	L	R	T	Q	V	N	T	Q	A	E	Q	L	R	R	Q	L	T	138	A-IV-8
	139	P	Y	A	Q	R	M	E	R	V	L	R	E	N	A	D	S	L	Q	A	S	L	R	160	A-IV-9
	161	P	H	A	D	E	L	K	A	K	I	D	Q	N	V	E	E	L	K	G	R	L	T	182	A-IV-10
	183	P	Y	A	D	E	F	K	V	K	I	D	Q	T	V	E	E	L	R	R	S	L	A	204	A-IV-11
	205	P	Y	A	Q	D	T	Q	E	K	L	N	H	Q	L	E	G	L	T	F	Q	M	K	226	A-IV-12
	227	K	N	A	E	E	L	K	A	R	I	S	A	S	A	E	E	L	R	Q	R	L	A	248	A-IV-13
	249	P	L	A	E	D	V	R	G	N	L	R	G	N	T	E	G	L	Q	K	S	L		269	A-IV-14
	270				A	E	L	G	G	H	L	D	Q	Q	V	E	E	F	R	R	R	V	E	288	A-IV-15
	289	P	Y	G	E	N	F	N	K	A	L	V	Q	Q	M	E	Q	L	R	T	K	L	G	310	A-IV-16
	311	P	H	A	G	D	V	E	G	H	L	S	F	L	E	K	D	L	R	D	K	V	N	332	A-IV-17
E	62	A	L	M	D	E	T	M	K	E	L	K	A	Y	K	S	E	L	E	E	Q	L	T	83	E-4
	84	P	V	A	E	E	T	R	A	R	L	S	K	E	L	Q	A	A	Q	A	R	L	G	105	E-5
	106												A	D	M	E	D	V	C	G	R	L	V	116	E-6
	117	Q	Y	R	G	E	V	Q	A	M	L	G	Q	S	T	E	E	L	R	V	R	L	A	138	E-7
	139	S	H	L	R	K	L	R	K	R	L	L	R	D	A	D	D	L	Q	K	R	L	A	160	E-8
	161	V	Y	Q	A	G	A	R	E	G	A	E	R	G	L	S	A	I	R	E	R	L	G	182	E-9
	183	P	L	V	E	Q	G	R	V	R	A		A	T	V	G	S	L	A	G	Q			201	E-10
	202	P	L	Q	E					R	A	Q	A	W	G	E	R	L	R	A	R			217	E-11
	218			M	E	E	M	G	S	R	T	R	D	R	L	D	E	V	K	E	Q	V	A	237	E-12
	241	A	K	L	E	E	Q	A	Q	Q	I	R	L	Q	A	E	A	F	Q	A	R	L	K	262	E-13
	267	P	L	V	E	D	M	Q	R	Q	W	A	G	L	V	E	K	V	Q	A	A	V	G	288	E-14
A-II	40												S	Y	F	E	K	S	K	E	Q	L	T	50	A-II-4
	51	P	L	I	K	K	A	G	T	E	L	V		N	F	L	S	Y	F	V	E	L	G	71	A-II-5
C-III	41												G	W	V	T	D	G	F	S	S	L	K	51	C-III-4
	52	D	Y	W	S	T	V	K	D	K	F	S		E	F	W	D	L	D	P	E	V	R	72	C-III-5
C-II	51	D	L	Y	S	K	S	T	A	A	M	S	T	Y	T	G	I	F	T	D	Q	V	L	72	C-II-4
C-I	40	E	W	F	S	E	T	F	Q	K	V	K												50	C-I-4
LALI	43												S	L	Y	T	D	T	V	A	V	L	T	53	LALI-4
	54	P	Y	L	Q	K	I	R	E	N	V	T	K	M	Y	Q	V	Y	V	E	S	K	Q	75	LALI-5

Fig. 3. Internal repeats in human and lamprey apolipoproteins. LAL1, a lamprey apolipoprotein isolated by Pontes et al. (176). a. *(Top):* The last 33 amino acids of exon 3 can be divided into three repeats of 11 amino acids. Colors (ref. 28) indicate proline (purple, P); aspartic acid or glutamic acid (red, D and E); arginine or lysine (blue, R and K); and methionine, valine, leucine, isoleucine, phenylalanine, tyrosine, or tryptophan (green, M, V, L, I, F, Y, and W). The remaining amino acids, glycine, alanine, serine, threonine, asparagine, glutamine, histidine, and cysteine (G, A, S, T, N, Q, H, and C) are uncolored and called indifferent. Any column containing 4 or more amino acids of a single color is said to possess that character and the amino acids of that color are boxed. Numbers along the left and right margins are residue numbers in the mature peptide. We refer to these three repeats as the first three repeats in each gene. b. *(Bottom):* Most of the repeats in exon 4 are 22-mers, each of which is made up of two 11-mers, and the other repeats are 11-mers. When more than 16 amino acids in a column are of a single color, the amino acids with that color are boxed. We refer to the first repeat in exon 4 as the fourth repeat in each gene; the repeats are denoted as A-I-4, A-I-5, etc.

tures described as the basic lipid-binding domains by Fukushima et al. (170) and Nakagawa et al. (172). Indeed, some of the repeats identified in Fig. 3 represent lipid-binding amphipathic helicies identified in the individual proteins (173). Luo et al. (27) have designated the three repeats of 11 codons in exon 3 the first three repeats in each

Downloaded from www.jlr.org by guest, on June 19, 2012

gene and the first 22 codons in exon 4 the fourth repeat, e.g., A-I-4 (Fig. 3 *bottom*).

In general, among the three longer proteins (A-I, A-IV, and E), the repeats in apoA-IV are the most regular, whereas those in apoE are the least regular. In fact, fewer of the repeats in apoE begin with proline, deletions appear to have occurred in repeats E-10, E-11, and E-12, and there are extra residues (not shown) between E-12 and E-13 and between E-13 and E-14.

Some remarkable features emerge from the alignment shown in Fig. 3 *bottom*. Most prolines (purple) appear in the first column of group A. Columns 6 and 10 of both groups A and B are mostly occupied by hydrophobic amino acids (green), the rest being indifferent (uncolored) amino acids. In both groups, columns 4 and 5 are predominantly occupied by acidic amino acids (red), and column 9 by basic amino acids (blue), while column 11 does not have a predominant pattern (color). Column 2 in group A is predominantly occupied by hydrophobic amino acids, but that in group B does not have a predominant pattern. Furthermore, hydrophobic amino acids (green) occur considerably more frequently in column 3 in group B than in column 3 in group A, and basic amino acids (blue) occur more frequently in column 7 in group B than in column 7 in group A. Thus, while the two groups have major similarities, they also have minor differences.

Complete cDNA sequences have recently become available for chicken apolipoproteins (174, 175) and for two lamprey apolipoproteins named LAL1 and LAL2 (176). It has been found that chicken apoA-I has a repeat pattern very similar to that in human apoA-I (174, 175), suggesting that all the internal repeats in these sequences arose before the bird-mammal split. Lamprey LAL1 has a repeat pattern similar to those in human apoA-II and C-III (Fig. 3). On the other hand, the sequence for LAL2 provides no clear indication for the existence of internal repeats. In fact, even the common block of 33 residues seen in apoA-I, A-II, A-IV, C-I, C-II, C-III, and E cannot be identified in LAL2. However, the segments defined by the three proline residues at positions 57, 79, and 100 in the mature peptide region of LAL2 could be remnants of three internal repeats. The segments 79-99 and 100-120 have the potential to form an amphipathic helical structure (data not shown), though the segment 59-78 does not have such a potential. The segment 35-56 can also form an amphipathic helical structure and might represent another repeat.

Under the pattern of internal repeats depicted in Fig. 3, the mature peptide of each of these proteins (except LAL2) is almost completely made up of internal repeats. For example, in apoA-II, A-I, and C-III, only the first and the last six or seven amino acids of the mature peptide are not included in the repeats shown in Fig. 3.

For the genes shown in Fig. 3 where the genomic structure is known, the last intron invariably interrupts the coding sequence between the second and third base of the last amino acid residue of the common block. The genomic structure of two of the apolipoproteins in this figure, apoC-I and LAL1, is still unknown. However, based on the alignment shown, we speculate that the last intron in apoC-I and LAL1 might interrupt these sequences at residues 39 and 42, respectively. As discussed in the section on the Structural organization of the apolipoprotein genes, secondary structural analysis of apoC-I also predicts the presence of an intron at the same location (residue 39). A similar analysis (data not shown) of the lamprey apolipoproteins LAL1 and LAL2 also suggests that the former might have an intron at the predicted location (residue 42), and the latter might have an intron at or around residue 34.

The similarities in genomic structure and in the pattern of internal repeats among the genes coding for the soluble apolipoproteins discussed above allow us to present a common model for the structure of their mRNAs (**Fig. 4**). Based on such a model, one can directly compare regions of the gene (and protein) among different apolipoproteins, as well as compare the same apolipoproteins (and their genes) among different species. From such sequence comparisons, one can estimate the rates of evolution of individual apolipoproteins and their subdomains, and infer the evolutionary relationships among the members of the apolipoprotein multigene family (see below).

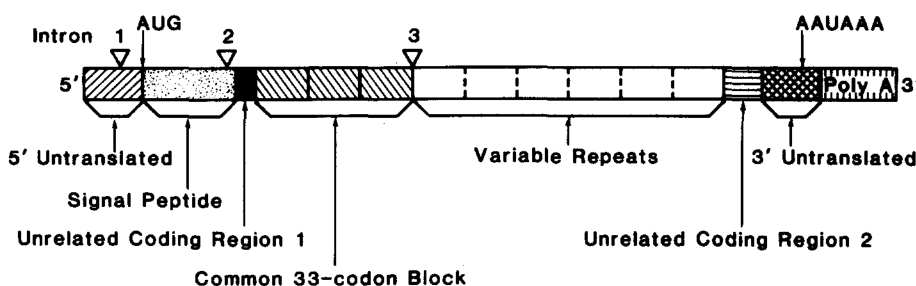

Fig. 4. Common structure of apolipoprotein mRNAs. The mRNAs that share this common structure include: apoA-I, A-II, C-I, C-II, C-III, E, and apoVLDL-II. ApoA-IV mRNA has a similar structure except that it has lost the intron 1 represented by the first triangle. The mRNA for LAL1, a lamprey apolipoprotein, also shares the same structure, but the location of the introns is unknown.

Downloaded from www.jlr.org by guest, on June 19, 2012
ASBMB JOURNAL OF LIPID RESEARCH

ApoB-100 structure and internal repeats

ApoB-100 is the largest of the apolipoproteins. For many years, investigations into the structure of apoB have been hampered by its large size, its insolubility in aqueous buffers after delipidation, and its susceptibility to proteolytic cleavage during purification. ApoB-48, the form of apoB present in chylomicrons, appears to share epitopes (88, 89) and sequences (89, 97) common to the NH_2-terminal half of apoB-100. Recently, the application of recombinant DNA techniques has allowed the cloning (114–116, 118, 120, 121, 123) and complete sequence analysis of apoB-100 cDNAs (65, 113, 117, 119, 122). The reactivity of some of the cloned cDNAs to specific monoclonal antibodies has allowed the assignment of apoB-74 to the carboxyl-terminal, and apoB-26, the amino-terminal parts of the molecule (121). The deduced amino acid sequence of apoB-100 has provided valuable information on some of the structural aspects of apoB function (65, 113, 117, 119, 122). ApoB-100 is a 4536-amino acid protein, the largest polypeptide ever studied (65, 113, 117, 119, 122). It is the most hydrophobic of all the apolipoproteins. The sequence is punctuated by 25 cysteines, 12 of which are located in the NH_2-terminal 500 residues. Inspection of the sequence has led to the identification of two potential LDL receptor-binding domains based on their homology to the receptor-binding region in apoE (116). A synthetic peptide (residues 3345–3381) of one of the two regions was found to bind to the LDL receptor and suppress 3-hydroxy-3-methylglutaryl coenzyme A reductase activities in cultured human fibroblasts (122). In addition, Yang et al. (122) have extensive sequence data from direct analysis of LDL apoB-100 tryptic peptides. Based on the differential accessibility of the different regions of LDL apoB-100 to tryptic digestion, they have proposed a model for the structure and conformation of apoB-100 on LDL (122).

Immunochemical polymorphism of human apoB has been shown in various laboratories (177–179). Recent studies (180) using monoclonal antibodies against purified LDL containing apoB-100 as its only protein component confirm the existence of such immunochemical polymorphisms. The first evidence for polymorphism at the primary amino acid sequence level was provided by Wei et al. (121) who found, scattered over the carboxyl-terminal fifth of the protein, differences in single amino acid residues identified by direct sequencing of apoB-100 tryptic peptides and those deduced from the cDNA sequence. It became evident later that such polymorphisms indeed existed throughout the molecule as revealed by complete cDNA sequences reported by different laboratories (65, 113, 117, 119, 122). Some of the differences are reflected by restriction fragment length polymorphisms involving the apoB gene (92–94), while the others are confirmed by direct sequence analysis of apoB-100 tryptic peptides (122). It is possible that some of the sequence heterogeneities of apoB-100 involve crucial parts of the molecule resulting in differences in function and metabolism of LDL. Indeed, evidence that a specific apoB-100 allele can result in familial hypercholesterolemia and premature atherosclerosis has been found in a strain of pigs described by Rapacz et al. (181). These investigators found that these pigs bear three immunologically defined lipoprotein-associated markers (allotypes), including one that defines a variant apoB in these animals. Future studies on these genetic markers, especially with respect to apoB structure and function, will be important for understanding apoB-LDL receptor interactions.

When human apoB-100 was digested with trypsin, the amino acid composition of the remaining parts of the protein showed little difference from that of the undigested protein (182, 183). This observation suggests that apoB-100 contains internal repeats. This is an interesting possibility because it might explain how such an exceptionally large protein has evolved and because internal repeats have been found in all the apolipoproteins discussed above. The recent studies by Yang et al. (122) and DeLoof et al. (184) have provided evidence for the existence of many internal repeats in apoB-100. Of particular interest was the discovery of the following two classes of repeats: amphipathic helical regions and hydrophobic proline-rich domains (184; Tables 4 and 5). These repeats were detected by a special iterative multiple alignment algorithm (184). The threshold score used was 11, a level such that the well-documented internal repeats of apoA-I, A-IV, or E appear clearly on the comparison matrices. Furthermore, the statistical significance of these repeats was confirmed by Monte-Carlo simulation tests and other tests (184).

The first class of internal repeats, together with the 22-residue consensus, is shown in **Table 4.** Representation of the consensus sequence in an Edmundson-wheel diagram indicates that it has an amphipathic structure: hydrophobic amino acids are located on one side of the helix, whereas the polar residues are located on the other side (**Fig. 5**). This is also true for each of the repeats in Table 4. Moreover, calculations of the mean hydrophobicity and the helical hydrophobic moment of each repeat and the consensus sequence indicate that all the sequences have a high helical hydrophobic moment consistent with an amphipathic structure. Such structures have also been found in other apolipoproteins and are thought to be important for phospholipid binding (see above). Interestingly, on intact LDL particles, the domains containing these repeats are generally inaccessible to trypsin (122), further supporting the hypothesis that these are involved in lipid binding.

In contrast, the proline-rich repeats (**Table 5**) are unique to apoB-100 and are characterized by the preponderance of hydrophobic residues. Their secondary structure is predicted to be composed of predominantly β-sheets and β-turns (due to proline residues), which might interact with lipids in a particular way (184). Computer modeling (Brasseur, R., H. De Loof, M. Rosseneu, and J. M. Ruysschaert, personal communication) of these proline-rich sequences in the presence of dipalmitoylphosphatidylcholine suggests

Downloaded from www.jlr.org by guest, on June 19, 2012
ASBMB
JOURNAL OF LIPID RESEARCH

TABLE 4. 22 Residue consensus

First Residue	Sequence	Score	Mean Hydrophobicity	Mean Helical Hydrophobic Moment
2079	QFVRKYRAALGKLPQQANDYLN	12.54	-.20	0.98
2135	DAKINFNEKLSQLQTYMIQFDQ	11.63	-.08	0.89
2173	NIIDEIIEKLKSLDEHYHIRVN	12.09	-.06	1.02
2384	TFIEDVNKFLDMLIKKLKSFDY	11.59	.06	1.01
2407	QFVDETNDKIREVTQRLNGEIQ	12.31	-.32	1.02
4150	RVTQEFHMKVKHLIDSLIDFLN	12.68	.03	1.00
4237	DVISMYRELLKDLSKEAQEVFK	11.95	-.12	0.99
4397	EYIVSASNFTSQLSSQVEQFLH	11.85	.13	0.69
4463	DYHQQFRYKLQDFSDQLSDYYE	12.90	-.31	0.82
Consensus	**DFIDEFNEKLKDLSDQLNDFLN**		-.13	0.98

The 22-residue-long consensus sequence derived by the iterative alignment procedure. This consensus sequence is shown in an Edmundson wheel (224) representation on Fig. 5. Identical residues are printed bold and related amino acids are underlined. The mean scores are calculated using the Staden (225) scoring matrix. The mean hydrophobicity and mean helical hydrophobic moment are calculated for all segments as previously described (159, 226). The mean hydrophobicity of the different segments is close to zero. The mean helical hydrophobic moment is, however, close to unity for most segments. This is indicative of the amphipathic nature of these segments when oriented in a helical conformation. The presentation of the individual segments in an Edmundson-wheel diagram confirms that each segment can form an amphipathic helix. This table has been taken from DeLoof et al. (184).

that the first part of each segment consists of a β-sheet that might penetrate into the acyl chains. After a turn around a proline residue, the segment can form a second β-sheet parallel to the first one, but with a reverse orientation. Such a structure would be able to penetrate more deeply into the LDL than the amphipathic helices. Cooperativity in the lipid-binding of these two different classes of subdomains, i.e., amphipathic versus proline-rich regions, might account for the observation that, in contrast to the smaller apolipoproteins, apoB does not exchange between different lipoprotein particles (3).

ApoD structure: lack of internal repeats

ApoD is a glycoprotein with an apparent M_r of 33,000 found in HDL (185, 186). Its function is unknown. However, the protein was found complexed with LCAT, and cholesteryl ester transport is one of its postulated functions (185, 187). The structure of apoD was recently deduced from its cloned cDNA sequence (66). It is a 169-amino acid protein with a predicted α-helical content (158) of less than 5%, in contrast to the substantial amounts of α-helix present in other apolipoproteins. We performed a dot matrix analysis of the human apoD sequence and have not uncovered any potential repeats. This sets apoD distinct from the other apolipoproteins sequenced to date.

EVOLUTION

One important aspect in the study of evolution of genes is to know the rate of nucleotide substitution. Because of the degeneracy of the genetic code, nucleotide substitutions in protein coding genes can be of two types: synonymous and nonsynonymous. A synonymous substitution causes no amino acid change, while a nonsynonymous substitution results in an amino acid replacement. Treating these two types of substitutions separately can give a better insight into the mechanisms of nucleotide substitution in evolution. Moreover, nucleotide sites can also be classified into synonymous and nonsynonymous. The conventional approach is to count a nucleotide site as a nonsynonymous site if all possible changes at that site are nonsynonymous, two-thirds nonsynonymous, and one-third synonymous if

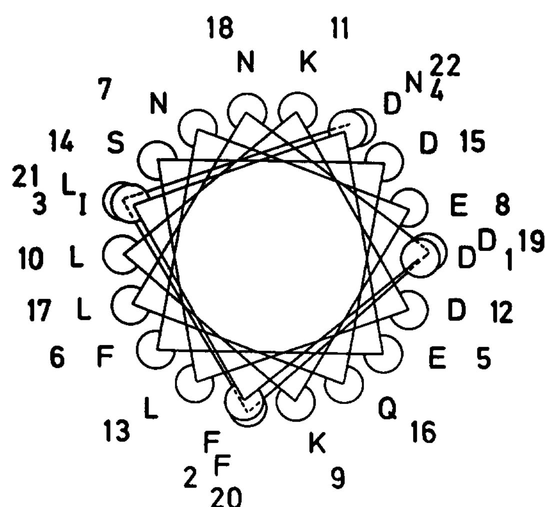

Fig. 5. Edmundson-wheel diagram of the 22-residue consensus sequence in apoB-100. The diagram is drawn according to Schiffer and Edmundson (224). One side of the helix clearly has hydrophobic amino acids while nonpolar amino acids are present on the other side of the helix. Edmundson-wheel diagrams of individual repeats in Table 2 show that each of them can potentially form an amphipathic helix. This figure is taken from DeLoof et al. (184).

Downloaded from www.jlr.org by guest, on June 19, 2012
ASBMB JOURNAL OF LIPID RESEARCH

TABLE 5. "Proline-rich" consensus sequences

First Residue	Sequence	Score
A. 52 Residues		
1283	LKMLETVRT PALH FKSVGFHLPS REFQVPTFT I PKLYQLQVP-LLGVLDLSTN	11.88
2574	EVS LQALQKATFQTPDF IVPLTDLRIPSVQINFKDLKNIKI PSRFSTPEFTI-	11.31
2666	LRDLKVED I PLARITLPDFRLPEIAIPEF I IPTLNLNDFQVP-DLHI PEFQLP	12.28
3245	YV F PKAVSMPSFS I LGSDVRVPSYTLILP SLELPVLHVPRNL-KLSLPDFKEL	11.98
3711	NDLNSVLVMPTFHVPFTDLQVPSCKLDFREIQIYKKLRTSS F-ALNLPTLPEV	11.40
3805	SDGIAALD LNAVANKIADFELPTI IVPEQTIEIPSIKFSVPA-GIAIPSFQAL	12.26
Consensus	**LD SLKALDMPTFHI PSSDFRLPS ITIPEPTIEIPKLKNSQVP-ALS I PDFQEL**	
B. 25 Residues		
1296	FKSVGFHLPSRE FQV PTFTI PKL YQ	11.84
2585	FQTPDF IVPLTD LRI PSVQI NFKDL	12.76
2629	FH I PSFTIDFVEMKVKI I RTI DQML	11.68
2704	FQVPDLHI PEFQL PHI SHTI E VPTF	12.72
3219	FQ I PGYTVPVVNVEVSPFTI EMSA F	12.80
3251	VSMPSF S I LGSDVRVPSYT L I LPSL	12.24
3717	LVMPTFHVPFTD LQVPSCKLDFRE I	12.54
3823	FE LPTI I VPEQT I E I PSIKFSVPAG	12.64
Consensus	**FQMPSFHVPETDLEVPS I T I EVPA L**	

Fifty-two- and 25-residue-long consensus sequences derived by the iterative alignment procedure. Identical residues are printed bold and related amino acids are underlined. The mean scores are calculated using the Staden (225) scoring matrix and gaps are penalized. This table has been taken from DeLoof et al. (184).

two of the three possible changes at that site are nonsynonymous, and so on (188). Following this classification, one can then calculate the number (K_A) of substitutions per nonsynonymous site and the number (K_S) of substitutions per synonymous site when comparing two homologous genes (188). When the divergence time between the two genes is known, then one can compute the synonymous and nonsynonymous rates of nucleotide substitution.

Knowing the rate of nucleotide substitution can serve at least two purposes. First, it may enable one to infer the dates of certain types of evolutionary events such as gene duplication. By this approach, rough estimates of the duplication dates have been obtained for several pairs of the apolipoprotein genes (27). Second, it may enable one to infer the stringency of structural requirements in a protein or parts of a protein. Comparative studies of protein and DNA sequences have led to the general conclusion that functionally more important molecules or parts of a molecule evolve more slowly than less important ones (188–191). A good example is the proinsulin polypeptide, which consists of three peptides, A, B, and C. The C peptide connects the A and B chains of the proinsulin and is cleaved off following disulfide linkage and assembly of the latter two chains to form the mature insulin molecule. The A and B chains comprise the bioactive insulin that binds to the insulin receptor, while the C peptide, once cleaved from proinsulin, does not have any known function. The rate of nonsynonymous substitution is found to be much higher in the region coding for the C peptide than in the regions coding for the A and B peptides, suggesting that the former is subject to less stringent structural constraints than the latter (191). Using this idea, Luo et al. (27) and Datta et al. (133) have inferred the structure–function relationship in most of the apolipoproteins.

Another important aspect in the study of evolution of genes is to infer the evolutionary relationships among genes. For the apolipoprotein genes, this can be done by considering similarities in DNA or amino acid sequences, pattern of internal repeats, and genomic structure (27; see also Fig. 4).

Rates of nucleotide substitution

Luo et al. (27) and Datta et al. (133) have studied the rate of nucleotide substitution in the genes coding for apoA-I, A-II, A-IV, C-II, C-III, and E. The results of their analysis are shown in **Table 6**. In this discussion, the rate of synonymous substitution is not considered for two reasons. First, this rate is not useful for inferring the stringency of structural requirements at the amino acid level. Second, the degree of divergence between different apolipoprotein genes is very large at synonymous sites so that it is difficult to obtain reliable estimates of the K_S values. Therefore, the K_S values cannot be used to infer the duplication dates

Downloaded from www.jlr.org by guest, on June 19, 2012

ASBMB JOURNAL OF LIPID RESEARCH

TABLE 6. Number of substitutions per nonsynonymous site between human and dog or rat apolipoprotein genes

Genes	Signal Peptide Region	Repeats in the Mature Peptide Region																Total Mature Peptide
		1+2+3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		
β-globin																		0.12 (0.02)
B-100																		0.09 (0.01)
C-II	0.13 (0.06)	0.26 (0.07)	0.11 (0.05)															0.22 (0.04)
C-III	0.06 (0.05)	0.10 (0.04)	0.08 (0.06)	0.23 (0.08)														0.17 (0.03)
	0.15 (0.08)	0.26 (0.07)	0.28 (0.13)	0.26 (0.08)														0.33 (0.05)
	0.09 (0.05)	0.23 (0.06)	0.34 (0.15)	0.31 (0.10)														0.32 (0.05)
A-II	0.19 (0.08)	0.25 (0.06)	0.29 (0.12)	0.51 (0.14)														0.35 (0.05)
A-I	0.15 (0.07)	0.08 (0.03)	0.25 (0.08)	0.13 (0.05)	0.23 (0.10)	0.19 (0.07)	0.49 (0.14)	0.43 (0.12)	0.51 (0.14)	0.21 (0.07)	0.14 (0.08)	0.46 (0.12)						0.25 (0.02)
E	0.22 (0.09)	0.08 (0.03)	0.19 (0.07)	0.07 (0.04)	0.23 (0.11)	0.21 (0.07)	0.07 (0.04)	0.10 (0.05)	0.48 (0.15)	0.27 (0.11)	0.26 (0.09)	0.12 (0.05)	0.14 (0.06)					0.19 (0.02)
A-IV	0.16 (0.08)	0.09 (0.04)	0.15 (0.06)	0.40 (0.13)	0.26 (0.14)	0.25 (0.08)	0.41 (0.12)	0.45 (0.14)	0.23 (0.08)	0.26 (0.09)	0.19 (0.07)	0.39 (0.12)	0.13 (0.06)	0.28 (0.10)	0.23 (0.08)	0.15 (0.06)		0.24 (0.02)

For apoC-II, the comparison is between human and dog; for apoC-III, the three comparisons are between human and dog, dog and rat, and human and rat, respectively; for the other genes, the comparison is between human and rat. The apoB-100 is a partial sequence. The values in parentheses are standard errors. The repeat numbers in different genes do not necessarily correspond to each other. The β-globin gene is given for comparison because it evolves at the average rate for 35 mammalian genes (188).

or the evolutionary relationships among the apolipoprotein genes. In the present section, consideration is given to the substitution rates in the signal peptide region and in the mature peptide region as a whole. The substitution rates in individual repeat regions are considered later.

In the signal peptide region, the K_A value for apoC-III is 0.06 between human and dog, 0.15 between dog and rat, and 0.09 between human and rat. Since the three species diverged at about the same time (192), we can show that the K_A value is $(0.06 + 0.09 - 0.15)/2 = 0.00$ in the human lineage, 0.06 in the dog lineage, and 0.09 in the rat lineage; for the method, see Fitch and Margoliash (193). This suggests that no nonsynonymous substitution has occurred in the signal peptide region in the human apoC-III gene since the time of mammalian radiation, i.e., about 80 million years ago (192). The variation in the K_A value among the other genes can be largely attributed to statistical fluctuations and so the signal peptides in the other proteins have evolved at about the same rate. Actually, in the divergence between dog and rat, the signal peptide in apoC-III has also evolved at a rate similar to those in the other proteins in Table 6. It is not clear why the signal peptide in human apoC-III has been so well conserved in evolution.

In the mature peptide region of the apoC-III gene, the K_A value is 0.17 between human and dog, 0.33 between dog and rat, and 0.32 between human and rat. Therefore, the K_A value is 0.08 in human lineage, 0.09 in the dog lineage, and 0.24 in the rat lineage. Thus, apoC-III has evolved about three times faster in the rat lineage than in the human and dog lineages. This difference in rate is considerably larger than Wu and Li's (194) estimate that the nonsynonymous rate is, on the average, 1.3 times higher in the rodent lineage than in the human lineage. It will be interesting to know whether any other apolipoproteins also evolve much faster in the rodent lineage than in the human lineage.

In the mature peptide region, the K_A value between the human and dog apoC-II genes is somewhat larger than that between the human and dog apoC-III genes, suggesting that apoC-II is somewhat less conservative than apoC-III. The K_A value between the human and rat apoC-III genes is close to that between the human and rat apoA-II genes. It is seen from the table that the apoA-II gene evolves two to three times faster than the β-globin gene, which evolves at the average rate for 35 mammalian genes (188). We may therefore conclude that apoA-II, apoC-II, and apoC-III all evolve very rapidly. Further, we note that apoA-I and apoA-IV have evolved twice as fast as the β-globin and that apoE has also evolved considerably faster than the β-globin. Thus, all of these apolipoproteins are not conservative in evolution. In contrast, the partial apoB-100 sequence corresponding to residues 595 → 979 seems to be

JOURNAL OF LIPID RESEARCH ASBMB

Downloaded from www.jlr.org by guest, on June 19, 2012

more conservative than the β-globin (Table 6). It will be interesting to know whether this is true for the entire apoB-100 molecule. It is also interesting to note that the shorter proteins A-II, C-II and C-III are less conservative than the longer proteins A-I, E, A-IV, and apoB-100.

Evolutionary relationships among apoA-I, A-II, A-IV, C-I, C-II, C-III, E, LAL1, and LAL2

As discussed above, the genes coding for apoA-I, A-II, C-I, C-II, C-III, and E each have a total of three introns at the same locations and the gene encoding apoA-IV also has the same genomic structure except that it has lost the first intron. All of these genes have a similar repeat pattern and their third exon shares a common block of 33 codons (see Table 3 and Figs. 3 and 4). Based on these observations, Luo et al. (27) and Boguski et al. (195) have proposed that these genes are evolutionarily related and arose from a common ancestral gene. The evolutionary scenario proposed by Luo et al. (27) for these genes is as follows (**Fig. 6**).

The common ancestor of the apolipoprotein genes was very similar to the present day apoC-I in structure and length. It contained three introns at the same locations as observed in the present day genes. The third exon of this gene contained the common block of 33 codons mentioned above and the fourth exon contained the basic repeat unit of 11 codons, as does the present day apoC-I. This gene was duplicated into two; one of them led to apoC-I and the other became the common ancestor of the other apolipoprotein genes. In the latter lineage, the first 11 codons of exon 4 were duplicated and then the whole gene was duplicated into two. One of the two genes became the present day apoC-II, and the other became the common ancestor of the apolipoprotein genes other than apoC-I and apoC-II. In the latter lineage, a duplication of the first 22 codons of exon 4 occurred and then a duplication of the whole gene followed. In one of the two resultant genes, a deletion of the first 11 codons of exon 4 occurred and the gene was later duplicated, one part leading to apoA-II and the other to apoC-III. The other gene experienced six or seven duplications of the 22 codon repeat and a duplication or deletion of 11 codons in exon 4 and was then duplicated: one of them gained two duplications of 22 codons and became the present day apoE; the other gained one duplication of 22 codons in exon 4 and was then duplicated, one leading to apoA-I and the other to apoA-IV after gaining three duplications of 22 codons in exon 4. After the separation of apoA-IV from apoA-I, the gene for apoA-IV lost the first intron. A tentative evolutionary history of the internal repeats in apoA-I has been proposed by Fitch, Smith, and Breslow (196).

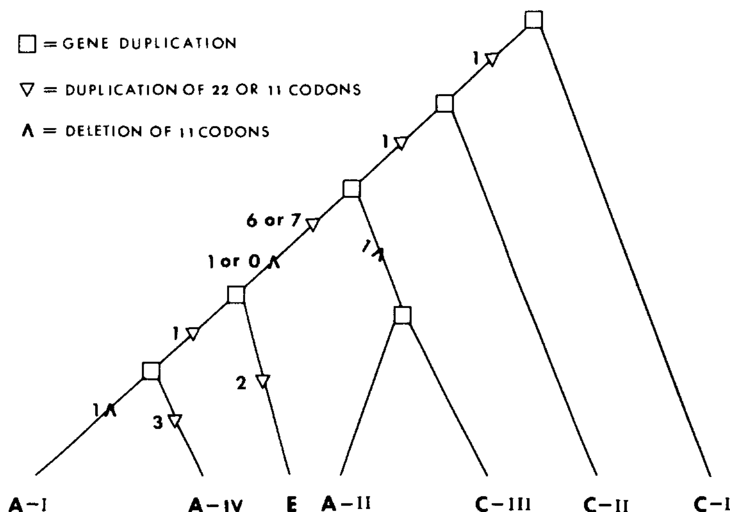

Fig. 6. A hypothetical scheme for the evolution of apolipoprotein genes. The primordial apolipoprotein gene was probably very similar to apoC-I in structure and length. It was duplicated in two; one of them led to apoC-I and the other became the common ancestor of the apolipoprotein genes. In the latter lineage, the first 11 codons of exon 4 were duplicated and then the whole gene was duplicated into two. One of them led to apoC-II and the other became the common ancestor of the apolipoprotein genes other than apoC-I and C-II. In the latter lineage, the first 22 codons of exon 4 were duplicated and a duplication of the whole gene followed. In one lineage, a deletion of the first 11 codons of exon 4 occurred and then the gene was duplicated into two, leading to apoA-II and C-III. In the other lineage, at least six duplications of 22 codons occurred (these six repeats are referred to as A-I-7, A-I-8, A-I-9, A-I-11, A-I-12, and A-I-13 in apoA-I, A-IV-7, A-IV-8, A-IV-9, A-IV-14, A-IV-16, and A-IV-17 in apoA-IV, and E-7, E-8, E-10, E-12, E-13, and E-14 in apoE) and possibly a duplication of 11 codons or a duplication of 22 codons followed by a deletion of 11 codons (this explains the half repeats A-I-6, A-IV-6, and E-6). A duplication of the whole gene then occurred. One of the two resultant genes led to apoE after gaining two duplications of 22 codons (repeats E-9 and E-11). The other gained one duplication of 22 codons (A-I-7 in apoA-I and A-IV-7 in apoA-IV) and was duplicated into two. One of them became apoA-IV after gaining three duplications of 22 codons (A-IV-12, A-IV-13 and A-IV-15). The other became apoA-I after losing 11 codons in repeat A-I-12. For more details, see Luo et al. (27).

Downloaded from www.jlr.org by guest, on June 19, 2012

Luo et al. (27) speculate that apoA-II and apoC-III are more closely related to each other than either of them is to the other genes (Fig. 6), because they are very similar in structure and length, e.g., their mature regions contain the same number of repeats and their signal peptide regions encoded by exon 2 differ by only one codon in length (Table 3 and Fig. 3). They also propose that apoA-I and apoA-IV are closer to each other than either is to apoE for two reasons [see also Boguski et al. (125)]. First, apoA-I and apoA-IV seem to share more repeats in common (Fig. 3). Actually, their repeat patterns are very similar, e.g., most of their repeats start with proline [see Fig. 3b and Boguski et al. (30)]. Second, the degree of sequence divergence is smaller between apoA-I and apoA-IV than between either of them and apoE (27). Furthermore, the evolutionary tree shown in Fig. 6 is the most parsimonious in terms of internal duplications and deletions. Barker and Dayhoff (26) had previously proposed that apoA-I separated from apoC-I earlier than did the ancestor of apoA-II and apoC-III. Under their proposal, one must assume that the repeats in exon 4 of apoA-II and apoC-III, and those in exon 4 of apoA-I were derived independently from the first 11 codons of exon 4 of apoC-I. This assumption does not seem plausible. It should be noted that the repeat A-II-5 and most of the repeats in apoA-I start with proline (Fig. 3b); therefore, it is quite possible that A-II-5 and most of the repeats in apoA-I share a common origin. At any rate, the tree proposed by Barker and Dayhoff (26) requires more internal duplication events than does that by Luo et al. (27).

The apolipoprotein genes shown in Fig. 6 are located on three separate human chromosomes: apoE, C-II, and C-I are on chromosome 19; apoA-I, C-III, and A-IV are on chromosome 11; and apoA-II is on chromosome 1 (see Table 2 and the section on Chromosomal localization and linkage analysis). The chromosomal distributions are not inconsistent with the evolutionary tree shown in Fig. 6 if we postulate the following scenario: a series of duplication events, starting from the ancestor of the apoC-I gene, produced on chromosome 19 the following apolipoprotein genes: E, A-I, C-III, A-II, C-II and C-I. Subsequently, the apoA-II gene moved to chromosome 1 and the genes for apoA-I and C-III moved to chromosome 11. The other apolipoprotein genes, i.e., those for E, C-I and C-II, remained as a cluster on chromosome 19. On chromosome 11, apoA-I underwent a duplication and produced apoA-IV. A DNA transposition then resulted in the insertion of the apoC-III gene in between the apoA-I and apoA-IV genes. This transposition is proposed for two reasons: first, nucleotide and amino acid sequence homology comparisons clearly indicate that apoA-I and apoA-IV are much more closely related to each other than either is to apoC-III. Second, since the apoC-III gene is in the opposite orientation to that of the apoA-I and apoA-IV genes (106, 136), it cannot be the immediate precursor of either of the latter two genes.

Using the numbers of nucleotide substitutions per nonsynonymous site between apolipoprotein genes and the rates of nonsynonymous substitution in apolipoprotein genes, Luo et al. (27) have estimated the divergence times between each pair of genes in Fig. 6 as follows. ApoA-I and apoA-IV diverged about 270 million years (Myr) ago, and their ancestor and apoE diverged about 420 Myr ago. ApoA-II and apoC-III separated from each other about 285 Myr ago, and their ancestor separated from that of apoA-I, apoA-IV, and apoE about 430 Myr ago. The last estimate is consistent with the observation that apoA-II is present in fish (176, 197-199) because fish and mammals diverged about 400 Myr ago (192). ApoC-II branched off about 570 Myr ago, and apoC-I and the other genes in Fig. 6 shared a common ancestor about 680 Myr ago. Luo et al. (27) cautioned that these are very rough estimates.

Lamprey LAL1 is similar to the mammalian apoA-II and apoC-III not only in repeat pattern (Fig. 3), but also in the length of mature peptide; human apoA-II and apoC-III are 77 and 79 residues long (Table 3) and lamprey LAL1 is 76 residues long (176). Further, like apoA-II, LAL1 appears to have a prosegment (176). Thus, LAL1 could be the counterpart of mammalian apoA-II. Since the estimated divergence date between A-II and C-III (285 Myr) is much more recent than the divergence date between the lamprey and mammalian lineages [about 400 to 450 Myr; see Romer (192)], it is reasonable to assume that LAL1 and the ancestor of A-II and C-III were derived from a common ancestral sequence at the time of the lamprey-mammal split. On the other hand, it is not clear whether the divergence between the ancestor of A-II and C-III and that of A-I, A-IV, and E (Fig. 6) occurred before or after the lamprey-mammal divergence for two reasons. First, the estimated date for the former divergence (430 Myr) is close to the latter divergence. Second, LAL1 is approximately equally similar to A-I, A-IV, E, A-II and C-III; for example, in the 33-residue common block, the numbers of identical amino acids between LAL1 and each of human A-I, A-II, A-IV, C-III, and E are 4, 5, 8, 5, and 5, respectively. It seems clear, however, that C-I and C-II had branched off earlier than the lamprey-mammal split because they differ from LAL1, A-II, and C-III in repeat pattern (Fig. 3). Lamprey LAL2 is very different from all known mammalian apolipoproteins and so its evolutionary relationship to the latter is uncertain.

Are apoB and apoD related to other apolipoproteins?

We have used the dot matrix method to search for potential homologous regions between apoB-100 and each of apoA-I, apoA-IV, and apoE. **Table 7** shows a partial list of regions of high similarity between apoE and apoB-100 and between apoA-IV and apoB-100. With one exception, all the similarities shown in Table 7 are equal to or higher than 30%. Although it is difficult to know whether any of

Downloaded from www.jlr.org by guest, on June 19, 2012

ASBMB JOURNAL OF LIPID RESEARCH JLR

TABLE 7. Percent similarities (S) between apoE and apoB-100 segments and between apoA-IV and apoB-100 segments

ApoE vs. ApoB-100				ApoA-IV vs. ApoB-100			
Segment in E	Segment in B-100	L[a]	S	Segment in A-IV	Segment in B-100	L[a]	S
21-67	1011-1057	47	23	77-104	323-350	28	39
42-67	2463-2488	26	35	104-135	381-412	32	31
47-75	2314-2342	29	31	110-133	2431-2454	24	33
48-86	316-354	39	31	148-173	1523-1548	26	35
66-86	2431-2451	21	38	202-226	1718-1742	25	32
121-156	745-780	36	39	217-237	1378-1398	21	33
122-143	464-485	22	36	232-260	1774-1802	29	38
133-155	3350-3372	23	39	224-254	561-591	31	32
156-177	3627-3648	22	36	243-264	1358-1379	22	36
203-223	1571-1591	21	38	269-289	1890-1910	21	38
209-228	2476-2495	20	35	331-363	2512-2544	33	30
231-252	2322-2343	22	41	332-353	1660-1681	22	30
233-261	2415-2443	29	31				
234-254	334-354	21	38				
238-262	4267-4291	25	32				
242-264	1092-1114	23	39				

[a]L, number of amino acid residues.

these segment pairs are true homologs, the results suggest that apoB-100 might be related to apoA-I, A-IV, and E [see also De Loof et al. (184)].

A dot matrix comparison between apoD (66) and the other apolipoproteins failed to show any significant homology. This suggests that evolutionarily apoD is not a member of the apolipoprotein multigene family comprised of the other apolipoproteins discussed in this review. Furthermore, analysis of RNA blots showed that apoD mRNA is most abundant in the adrenal gland, and is present in considerably higher concentrations in the kidney, pancreas, and small intestine than in the liver (66). Such a tissue distribution is quite distinct from that of all the other apolipoproteins and suggests that apoD might belong to a different class of proteins. The recent availability of apoD genomic structure supports such an interpretation (124). Unlike the other soluble apolipoproteins, the mature peptide coding portion of the gene contains three introns instead of one. Furthermore, the signal peptide region is not interrupted by any intron. The structural organization of the apoD gene is thus drastically different from those of the other soluble apolipoproteins. Indeed, the protein has been shown to display a high degree of homology to human plasma retinol-binding protein, α_1-microglobulin, ungulate β-microglobulin, rodent α_{2u}-globulin, and tobacco hornworm insecticyanin—all members of the α_u-globulin superfamily (66, 124).

Similarity of apolipoproteins to subdomains of LCAT

Both apoA-I and apoE can activate LCAT (200, 201). It is thus intriguing that we detected some similarities in the primary structure of both apoA-I and apoE to that of LCAT. **Table 8** presents the result of our comparison of apoA-I and apoE with LCAT. It is evident that both apolipoproteins have regions with a high similarity to regions in LCAT. We are, however, not certain whether any of the similarities represent true homologies, i.e., share a common origin.

RELATIVE RATES OF EVOLUTION AND FUNCTIONAL ASPECTS OF APOLIPOPROTEIN STRUCTURE

Luo et al. (27) and Datta et al. (133) have used the rate of nonsynonymous substitution to infer the stringency of structural requirements in various regions of the apolipoproteins. Their results are as follows.

TABLE 8. Percent similarities (S) between apoA-I and LCAT segments and between apoE and LCAT segments

ApoA-I vs. LCAT				ApoE vs. LCAT			
Segment in ApoA-I	Segment in LCAT	L[a]	S	Segment in ApoE	Segment in LCAT	L[a]	S
16-30	334-348	15	48	15-45	135-165	31	29
42-72	221-251	31	29	96-142	155-201	47	26
123-183	351-411	61	21	150-202	280-332	53	23
134-213	121-200	80	20	182-194	338-350	13	46
163-174	117-128	12	50	206-231	140-165	26	31
167-188	114-135	22	32	274-288	158-172	15	60

[a]L, number of amino acid residues.

Downloaded from www.jlr.org by guest, on June 19, 2012

ASBMB JOURNAL OF LIPID RESEARCH

As mentioned above, all of the apolipoprotein genes shown in Table 6 and Figs. 3 and 4 contain a common block of 33 codons in the mature peptide region. This block is located at the end of exon 3 and is referred to as repeats 1, 2, and 3 in Table 6. In this block, the K_A value (the number of nucleotide substitutions per nonsynonymous site) is low for apoA-I, A-IV, and E, suggesting that this block of 33 amino acids may be structurally important for the function of these apolipoproteins. This block is also well conserved in human and dog apoC-III, though not in apoC-II, apoA-II, and rat apoC-III.

The relatively high overall rates of nonsynonymous substitution in apoA-I, apoA-IV, and apoE suggest that the structural requirement of lipid binding is not stringent, for all three proteins contain many lipid-binding domains (30, 164, 173, 202). For example, each of the repeats A-I-8, A-I-9, A-I-10, A-I-13, A-IV-7, A-IV-10, E-11, and E-12 contains an amphipathic helix (106, 164, 173, 202); yet, they all exhibit high rates of nonsynonymous substitution.

ApoC-II

In apoC-II, exon 4 (repeat 4) is well conserved, probably because this part is important for the activation of lipoprotein lipase and in chylomicron and VLDL metabolism. Support for the importance of the carboxyl-terminal part of the molecule in lipoprotein lipase activation comes from the structure of apoC-$II_{Toronto}$, a mutant apoC-II isolated from a patient with homozygous deficiency (203). The sequence of apoC-$II_{Toronto}$ is identical to that of normal apoC-II from residues 1 → 68. Residues 69 → 79 are missing, being replaced by an unrelated hexapeptide, changes consistent with the deletion of a nucleotide for the codon of either Thr-68 or Asp-69 and a translation reading frame shift. ApoC-$II_{Toronto}$ is totally nonfunctional (203), indicating the importance of residues 69 → 79 for the normal lipase activating activity of the protein. This conclusion is consistent with previous experimental observations (203, 204). Studies using proteolytic fragments as well as synthetic peptides of human apoC-II indicate that the site of interaction of apoC-II with lipoprotein lipase is contained within residues 56–79 which encompasses repeat 4. Additional residues upstream (50–55 and 44–55) seem to enhance the enzyme activation in the presence of phospholipid-stabilized triolein emulsion because of the phospholipid binding activity that these additional residues confer to the peptide. It is interesting that Kinnunen et al. (204) and Smith et al. (205) found the last three residues (77–79) to be absolutely required for lipase activation; deletion of these residues resulted in total inactivation of the protein. Datta et al. (133) note that, in the dog, two of the last three amino acid residues are not conserved (-Gly-Asp-Ser instead of Gly-Glu-Glu), suggesting that the structural requirement for the terminal dipeptide -Glu-Glu is not absolute. Perhaps the critical requirement in the carboxyl-terminal tripeptide is the carboxylic acid side chains at the penultimate residue.

ApoC-III

In apoC-III, repeats 1, 2, and 3, and exon 4 are not well conserved. This observation suggests that the function of apoC-III (e.g., the modulation of hepatic uptake of lipoprotein remnant particles) does not have a very stringent structural requirement.

ApoA-II

The extremely high K_A values in all regions in apoA-II are probably due to weak functional constraints. ApoA-II does not seem to play an important functional role, although its exact function is unknown. In fact, apoA-II is either absent or expressed at a very low level in the dog (206), pig (207), cow (208), or chicken (209).

ApoA-I

A major function of apoA-I is the activation of LCAT (15, 18), an important enzyme in lipoprotein metabolism, which converts unesterified cholesterol to its ester form in plasma (18). Soutar et al. (16) have shown that both the amino- and carboxyl-terminal cyanogen bromide fragments (residues 1–85 and 147–243 of apoA-I) activate LCAT; in the latter fragment, residues 145–182 seem to be involved in the activation process (210). Surprisingly, this part of apoA-I sequence, which corresponds to repeats A-I-9 and A-I-10, is less conserved than the other repeats (Table 6). Thus, the structural requirements for LCAT activation may not be stringent. This conclusion is supported by other lines of evidence: *1)* another region of apoA-I also activates LCAT (see above); *2)* at least three other apolipoproteins, apoC-I (16), apoA-IV (17), and apoE (200, 201) can also activate LCAT; and *3)* synthetic model peptides that mimic apoA-I surface properties but differ from apoA-I in primary sequences are effective in LCAT activation (210–212). While LCAT activation may be a major function of apoA-I, a high rate of evolution may still occur since the protein can undergo considerable change in its primary structure without impairment of this function.

ApoA-IV

The function of apoA-IV is unknown, though the protein also has LCAT-activating potential (17). However, like apoA-I, although some of the regions in apoA-IV are relatively well conserved, most of them are not conserved. On the whole, apoA-IV evolves at a rate very similar to that in apoA-I. Thus, the function of apoA-IV also does not seem to have stringent structural constraints.

ApoE

ApoE is an important determinant in the interaction between apoE-containing lipoproteins and cell-surface receptors (213). Studies with specific proteolytic and chemical cleavage fragments of apoE indicate that the receptor-binding domain exists in the region between residues 129

Downloaded from www.jlr.org by guest, on June 19, 2012

ASBMB JOURNAL OF LIPID RESEARCH

and 191 (214). The use of monoclonal antibodies has localized the receptor-binding region to the vicinity of residues 140–150 (215). Studies of apoE mutants of known structure have further indicated the importance of residues 142, 145, and 146 in mediating receptor binding (216–218); the positive charge of arginine 158 also appears to be important in maintaining the correct conformation necessary for normal binding (219). Examination of the degree of conservation of the various repeats in apoE (Table 6) indicates that E-8 (residues 139–166), which encompasses the receptor-binding region, is indeed one of the most highly conserved sequences in apoE.

ApoB-100

ApoB-100 is an obligatory constituent of VLDL, IDL, and LDL and is the ligand that binds to the LDL receptor. It thus plays a central role in cholesterol homeostasis; in fact, the plasma concentration of apoB-100 is strongly correlated with the development of atherosclerosis (220, 221). The protein is unique among the apolipoproteins in its remarkable size and in its hydrophobicity and high β-sheet content (65, 113, 117, 119, 122). As discussed previously, apoB-100 appears to contain numerous internal repeated sequences. With such a structural organization, one would expect a considerable redundancy in apoB-100 structure and a relatively high rate of evolution. Interestingly, our analysis of the rate of nonsynonymous substitution in a substantial portion [about 1.2 kilobases (222)] of the apoB-100 gene indicates that it is the most conservative of all the apolipoprotein genes. The low rate of nonsynonymous substitution is all the more remarkable in that the sequence analyzed corresponds to residues 595 → 979, which does not overlap the putative receptor binding domains (116, 117, 122). This observation suggests that the apparent built-in redundancy in apoB-100 structure might be important in its overall function. It further supports the suggestion that apoB is a very ancient gene. An apoB-like protein, i.e., an LDL protein with an apparent M_r of $>250,000$ daltons and remarkably similar amino acid composition that is insoluble in aqueous media following delipidation, has been found in a number of fish species, including members of the lowest vertebrates, the hagfish, and the shark (197, 198, 223). The very similar properties of fish apoB-100 and human apoB-100 are also consistent with a low rate of evolution for this protein. However, to confirm whether apoB-100 is truly a well-conserved protein, we need additional sequences from the rest of the apoB-100 molecule from the rat as well as sequences of the protein from other species.

CONCLUDING REMARKS

In this review, we have presented our current knowledge of the biosynthesis, structure, and evolution of the apolipoprotein multigene family. The structure–function relationships of a number of apolipoproteins have also been examined from an evolutionary perspective. We mention below a number of problems that need to be explored further. First, the exact evolutionary relationship among the various apolipoproteins has to be constantly re-examined and updated as more data become available. For example, our estimates of the dates of the gene duplication events depicted in Fig. 6 have to be revised as more apolipoprotein genes are sequenced. Furthermore, the relationship of apoB to the other apolipoproteins is still uncertain, though we favor the hypothesis that they share a common ancestor in evolution. Additional sequence determinations from other animals such as birds and fish will be helpful in this respect. There is recent evidence that apoB-48 is the product of an intestinal mRNA that has an organ-specific in-frame stop codon. The exact mechanism by which the stop codon is introduced into the mRNA is currently under investigation. Second, we speculated on the evolutionary origin of the apoA-I/C-III/A-IV gene cluster in man (27). There is new evidence (Karathanasis, S. K., personal communication) that a similar relationship among the three apolipoprotein genes exists in chicken, a species that branched off about 270 Myr ago. This suggests that if indeed there had been a DNA transposition, it must have occurred prior to the establishment of the avian lineage. Future experiments on the structural organization of this gene cluster in more primitive animals will be quite informative in this respect. Third, we suggested that apoB-100 contains numerous internal repeats. One way to see whether this is true or not is to have sequence data from a bird species or a fish species. As discussed in the last section, additional apoB sequences from other species will also be useful in further defining the rate of evolution of apoB as a whole and of its various subdomains. The structure–function relationships among the latter can also be examined. Fourth, our observation of a threefold faster rate of evolution in apoC-III in the rat lineage than in the human and dog lineages has to be extended to other apolipoproteins in order to determine whether this is a general phenomenon. Fifth, the rate of evolution and the structure–function relationships in apoD should be studied when sequence data from one or more additional species become available. Sixth, our analysis and speculation on the structural constraints on the functional domains of individual apolipoproteins call for additional experiments that will confirm or modify our hypothesis. For example, previous experiments indicate the importance of the carboxyl-terminal tripeptide in apoC-II action, whereas we observed that at least two of the three residues can be changed during evolution. It is quite possible that the length rather than the exact amino acid sequence of this region of the molecule is important for LPL activation. This hypothesis can be tested in vitro.

With the recent availability of the human apoB-100 sequence, and the numerous unanswered questions in apolipoprotein structure, function, and evolution, we expect

Downloaded from www.jlr.org by guest, on June 19, 2012

ASBMB JOURNAL OF LIPID RESEARCH JLR

that the next decade will continue to be a rapidly growing one for lipoprotein research. Surely, new questions will arise as our knowledge increases in this area. But that is the gist of basic and clinical investigation!

We thank Ms. Suzanne Mascola for expert secretarial assistance, Bob Schwartz for drawing Fig. 3, Leslie Liu for assistance with the computer analysis of apolipoprotein secondary structure, and Drs. H. DeLoof and M. Rosseneu for collaboration on the analysis of apoB-100 internal repeats and sequence homology to other apolipoproteins reviewed in this paper (ref. 184). This work was supported by grants from the National Institutes of Health GM-30998 (to W.H.L.) and HL-16512 and HL-27341 (to L.C.), and a grant from the March of Dimes Birth Defects Foundation.
Manuscript received 8 April 1987 and in revised form 14 September 1987.

REFERENCES

1. Goldstein, J. L., and M. S. Brown. 1977. The low density lipoprotein pathway and its relation to atherosclerosis. *Annu. Rev. Biochem.* **46:** 897-930.
2. Kane, J. P., D. A. Hardman, and H. E. Paulus. 1980. Heterogeneity of apolipoprotein B: isolation of a new species from human chylomicron. *Proc. Natl. Acad. Sci. USA.* **77:** 2465-2469.
3. Kane, J. P. 1983. Apoprotein B: structural and metabolic heterogeneity. *Annu. Rev. Physiol.* **45:** 637-650.
4. Glickman, R. M., M. Rogers, and J. N. Glickman. 1986. Apolipoprotein B synthesis by human liver and intestine in vitro. *Proc. Natl. Acad. Sci. USA.* **83:** 5296-5300.
5. Brown, M. S., and J. L. Goldstein. 1986. A receptor-mediated pathway for cholesterol homeostasis. *Science.* **232:** 34-47.
6. Innerarity, T. L., and R. W. Mahley. 1978. Enhanced binding by cultured human fibroblasts of apoE-containing lipoproteins as compared with low density lipoproteins. *Biochemistry.* **17:** 1440-1447.
7. Sherrill, B. C., T. L. Innerarity, and R. W. Mahley. 1980. Rapid hepatic clearance of the canine lipoproteins containing only the E apoprotein by high affinity receptor. *J. Biol. Chem.* **255:** 1804-1807.
8. Hui, D. Y., T. L. Innerarity, and R. W. Mahley. 1981. Lipoprotein binding to canine hepatic membranes: metabolically distinct apoE and apoB,E receptors. *J. Biol. Chem.* **256:** 5646-5655.
9. LaRosa, J. C., R. I. Levy, P. Herbert, S. Z. Lux, and D. S. Frederickson. 1970. A specific apoprotein activator for lipoprotein lipase. *Biochem. Biophys. Res. Commun.* **41:** 57-62.
10. Breckenridge, W. C., J. A. Little, G. Steiner, A. Chow, and M. Poapst. 1978. Hypertriglyceridemia associated with deficiency of apolipoprotein C-II. *N. Engl. J. Med.* **298:** 1265-1273.
11. Brown, W-V., and M. L. Baginsky. 1972. Inhibition of lipoprotein lipase by an apolipoprotein of human very low density lipoprotein. *Biochem. Biophys. Res. Commun.* **46:** 375-382.
12. Wang, C. S., W. J. McConathy, H. U. Kloer, and P. Alaupovic. 1985. Modulation of lipoprotein lipase activity by apolipoproteins. Effect of apolipoprotein C-III. *J. Clin. Invest.* **75:** 384-390.
13. Windler, E., Y. S. Chao, and R. J. Havel. 1980. Determinants of hepatic uptake of triglyceride-rich lipoproteins and their remnants in the rat. *J. Biol. Chem.* **255:** 5475-5480.
14. Quarfordt, S. H., G. Michalopoulos, and B. Schirmer. 1982. The effect of human apolipoproteins on the in vitro hepatic metabolism of triglyceride emulsions in the rat. *J. Biol. Chem.* **257:** 14642-14647.
15. Fielding, C. J., V. G. Shore, and P. E. Fielding. 1972. A protein cofactor of lecithin:cholesterol acyltransferase. *Biochem. Biophys. Res. Commun.* **46:** 1493-1498.
16. Soutar, A. K., C. W. Garner, H. N. Baker, J. T. Sparrow, R. L. Jackson, A. M. Gotto, Jr., and L. C. Smith. 1975. Effect of the human apolipoproteins and phosphatidylcholine acyl donor on the activity of lecithin:cholesterol acyltransferase. *Biochemistry.* **14:** 3057-3064.
17. Steinmetz, A., and G. Utermann. 1985. Activation of lecithin:cholesterol acyltransferase by human apolipoprotein A-IV. *J. Biol. Chem.* **260:** 2258-2264.
18. Glomset, J. A. 1968. The plasma lecithin:cholesterol acyltransferase reaction. *J. Lipid Res.* **9:** 155-167.
19. Chan, L., R. L. Jackson, B. W. O'Malley, and A. R. Means. 1976. Synthesis of very low density lipoproteins in the cockerel: effects of estrogen. *J. Clin. Invest.* **58:** 368-379.
20. Chan, L., W. A. Bradley, R. L. Jackson, and A. R. Means. 1980. Lipoprotein synthesis in the cockerel liver: effects of estrogen on hepatic polysomal messenger RNA activities for the major apoproteins in very low density and high density lipoproteins and albumin, and evidence for precursors to these secretory proteins. *Endocrinology.* **106:** 275-283.
21. Chan, L. 1983. Hormonal control of apolipoprotein synthesis. *Annu. Rev. Physiol.* **45:** 615-623.
22. Gordon, J. I., D. P. Smith, E. Andy, D. H. Alpers, G. Schonfeld, and A. W. Strauss. 1982. The primary translation product of rat intestinal apolipoprotein A-I mRNA is an unusual preproprotein. *J. Biol. Chem.* **257:** 971-978.
23. Gordon, J. I., H. F. Sims, S. R. Lentz, C. Edelstein, and A. M. Scanu. 1983. Proteolytic processing of human preproapolipoprotein A-I. A proposed defect in the conversion of proA-I to A-I in Tangier's disease. *J. Biol. Chem.* **258:** 4037-4044.
24. Gordon, J. I., K. A. Budelier, H. E. Sims, C. Edelstein, A. M. Scanu, and A. W. Strauss. 1983. Biosynthesis of human preproapolipoprotein A-II. *J. Biol. Chem.* **258:** 14054-14059.
25. Gordon, J. I., C. L. Bisgaier, H. F. Sims, O. P. Sachder, R. M. Glickman, and A. W. Strauss. 1984. Biosynthesis of human apolipoprotein A-IV. *J. Biol. Chem.* **259:** 468-474.
26. Barker, W. C., and M. O. Dayhoff. 1977. Evolution of lipoproteins deduced from protein sequence data. *Comp. Biochem. Physiol.* **576:** 309-315.
27. Luo, C-C., W-H. Li, M. N. Moore, and L. Chan. 1986. Structure and evolution of the apolipoprotein multigene family. *J. Mol. Biol.* **187:** 325-340.
28. Fitch, W. M. 1977. Phylogenetics constrained by the crossover process as illustrated by human hemoglobins and a thirteen cycle, eleven amino acid repeat in human apolipoprotein A-I. *Genetics.* **86:** 623-644.
29. Karathanasis, S. K., V. I. Zannis, and J. L. Breslow. 1983. Isolation and characterization of the human apolipoprotein A-I gene. *Proc. Natl. Acad. Sci. USA.* **80:** 6147-6151.
30. Boguski, M. S., N. Elshourbagy, J. M. Taylor, and J. I. Gordon. 1984. Rat apolipoprotein A-IV contains 13 tandem repetitions of a 22 amino acid segment with amphipathic helical potential. *Proc. Natl. Acad. Sci. USA.* **81:** 5021-5025.
31. Boguski, M. S., M. Freeman, N. A. Elshourbagy, J. M. Taylor, and J. I. Gordon. 1986. On computer-assisted analysis of biological sequences: proline punctuation, consensus

ASBMB JOURNAL OF LIPID RESEARCH

Downloaded from www.jlr.org by guest, on June 19, 2012

sequences, and apolipoprotein repeats. *J. Lipid Res.* **27:** 1011-1034.
32. Wu, A-L., and H. G. Windmueller. 1979. Relative contributions by liver and intestine to individual plasma apolipoproteins. *J. Biol. Chem.* **254:** 7316-7322.
33. Radding, C. M., J. H. Bragdon, and D. Steinberg. 1958. The synthesis of low- and high-density lipoproteins by rat liver in vitro. *Biochim. Biophys. Acta.* **30:** 443-444.
34. Radding, C. M., and D. Steinberg. 1960. Studies on the synthesis and secretion of serum lipoproteins by rat liver slices. *J. Clin. Invest.* **39:** 1560-1569.
35. Marsh, J. B., and A. F. Whereat. 1959. The synthesis of plasma lipoprotein by liver. *J. Biol. Chem.* **234:** 3196-3200.
36. Hatch, F. T., Y. Aso, L. M. Hagopian, and J. Rubenstein. 1966. Biosynthesis of lipoprotein by rat intestinal mucosa. *J. Biol. Chem.* **241:** 1655-1665.
37. Roheim, P. S., L. I. Gidez, and H. A. Eder. 1966. Extrahepatic synthesis of lipoproteins of plasma and chyle: role of the intestine. *J. Clin. Invest.* **45:** 297-300.
38. Kessler, J. I., J. Stein, D. Dannacker, and P. Narcessian. 1970. Biosynthesis of low density lipoprotein by cell-free preparations of rat intestinal mucosa. *J. Biol. Chem.* **245:** 5281-5288.
39. Faloona, G. R., B. N. Stewart, and M. Fried. 1968. The effects of actinomycin D on the biosynthesis of plasma lipoproteins. *Biochemistry.* **7:** 720-725.
40. Chaikoff, I. L., S. Lindsay, F. W. Lorenz, and C. Entenman. 1948. Production of atheromatosis in the aorta of the bird by the administration of diethylstilbestrol. *J. Exp. Med.* **88:** 373-388.
41. Hillyard, L. A., C. Entenman, and I. L. Chaikoff. 1950. Concentration and composition of serum lipoproteins of cholesterol-fed and stilbestrol-injected birds. *J. Biol. Chem.* **223:** 348-359.
42. Katz, L. N., and L. R. Pick. 1961. Experimental atherosclerosis observed in the chicken. *J. Atheroscler. Res.* **1:** 93-100.
43. Luskey, K. L., M. S. Brown, and J. L. Goldstein. 1974. Stimulation of the synthesis of very low density lipoproteins in rooster liver by estradiol. *J. Biol. Chem.* **249:** 5939-5947.
44. Kudzma, D. J., P. M. Hegstad, and R. E. Stoll. 1973. The chick as a laboratory model for the study of estrogen-induced hyperlipidemia. *Metab. Clin. Exp.* **22:** 423-434.
45. Williams, D. L. 1979. Apoproteins of avian very low density lipoprotein: demonstration of a single high molecular weight apoprotein. *Biochemistry.* **18:** 1056-1063.
46. Jackson, R. L., H-Y. Lin, L. Chan, and A. R. Means. 1977. Amino acid sequence of a major apoprotein from hen plasma very low density lipoproteins. *J. Biol. Chem.* **252:** 250-252.
47. Chan, L., W. A. Bradley, and A. R. Means. 1980. Amino acid sequence of the signal peptide of apoVLDL-II, a major apoprotein in avian very low density lipoproteins. *J. Biol. Chem.* **255:** 10060-10063.
48. Blobel, G., and B. Dobberstein. 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immuoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J. Cell Biol.* **67:** 835-851.
49. Blobel, G., and B. Dobberstein. 1975. Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. *J. Cell Biol.* **67:** 852-862.
50. Chan, L., and W. A. Bradley. 1982. Signal peptides: properties and interactions. *In* Cellular Regulation of Secretion and Release. P. M. Conn, editor. Academic Press, New York. 301-321.
51. von Heijne, G. 1985. Signal sequences: the limits of variation. *J. Mol. Biol.* **184:** 99-105.
52. Walter, P., and V. R. Lingappa. 1986. Mechanism of protein translocation across the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* **2:** 499-516.
53. Blobel, G. 1980. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. USA.* **77:** 1496-1500.
54. Wickner, W. 1979. The assembly of proteins into biological membranes: the membrane trigger hypothesis. *Annu. Rev. Biochem.* **48:** 23-45.
55. Cheung, P., and L. Chan. 1983. Nucleotide sequence of cloned cDNA of human apolipoprotein A-I. *Nucleic Acids Res.* **11:** 3703-3715.
56. Shoulders, C. C., N. B. Myant, A. Sidoli, J. C. Rodriguez, C. Cortese, F. E. Baralle, and R. Cortese. 1985. Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than one gene per haploid genome. *Atherosclerosis.* **58:** 277-289.
57. Moore, M. N., F. T. Kao, Y. K. Tsao, and L. Chan. 1984. Human apolipoprotein A-II: nucleotide sequence of a cloned cDNA, and localization of its structural gene on human chromosome 1. *Biochem. Biophys. Res. Commun.* **123:** 1-7.
58. Knott, T. J., L. M. Priestley, M. Urdea, and J. Scott. 1984. Isolation and characterization of a cDNA encoding the precursor for human apolipoprotein A-II. *Biochem. Biophys. Res. Commun.* **120:** 734-740.
59. McLean, J. W., C. Fukazawa, and J. M. Taylor. 1983. Rat apolipoprotein E mRNA. Cloning and sequencing of double-stranded cDNA. *J. Biol. Chem.* **258:** 8993-9000.
60. McLean, J. W., N. A. Elshourbagy, D. J. Chang, R. W. Mahley, and J. M. Taylor. 1984. Human apolipoprotein E mRNA: cDNA cloning and nucleotide sequencing of a new variant. *J. Biol. Chem.* **259:** 6498-6504.
61. Knott, T. J., M. E. Robertson, L. M. Priestly, M. Urdea, and J. Scott. 1984. Characterization of mRNAs encoding the precursor of human apolipoprotein C-I. *Nucleic Acids Res.* **12:** 3909-3915.
62. Sharpe, C. R., A. Sidoli, C. S. Shelley, M. A. Lucero, C. C. Shoulders, and F. E. Baralle. 1984. Human apolipoproteins A-I, A-II, C-II, and C-III cDNA sequences and mRNA abundance. *Nucleic Acids Res.* **12:** 3917-3932.
63. Protter, A. A., B. Levy-Wilson, J. Miller, G. Bencen, T. White, and J. J. Seilhamer. 1984. Isolation and sequence analysis of the human apolipoprotein C-III gene and the intergenic region between apoA-I and apoC-III genes. *DNA.* **3:** 449-456.
64. Protter, A. A., D. A. Hardman, J. W. Schilling, J. Miller, V. Appleby, G. C. Chen, S. W. Kirsher, G. McEnroe, and J. P. Kane. 1986. Isolation of a cDNA clone encoding the amino-terminal region of human apolipoprotein B. *Proc. Natl. Acad. Sci. USA.* **83:** 1467-1471.
65. Chen, S. H., C-Y. Yang, P. F. Chen, D. Setzer, M. Tanimura, W-H. Li, A. M. Gotto, Jr., and L. Chan. 1986. The complete cDNA and amino acid sequence of human apolipoprotein B-100. *J. Biol. Chem.* **261:** 12918-12921.
66. Drayna, D., C. Fielding, J. McLean, B. Baer, G. Castro, E. Chen, L. Comstock, W. Henzel, W. Kohr, L. Rhee, W. Wion, and R. Lawn. 1986. Cloning and expression of human apolipoprotein D cDNA. *J. Biol. Chem.* **261:** 16535-16539.
67. Stoffel, W., E. Kruger, and R. Deutzmann. 1983. Cell-free translation of human liver apolipoprotein A-I and A-II mRNA. Processing of primary translation products. *Hoppe-Seyler's Z. Physiol. Chem.* **364:** 227-237.
68. Shoulders, C. C., A. R. Kornblihtt, B. S. Munro, and F. E. Baralle. 1983. Gene structure of human apolipoprotein A-I. *Nucleic Acids Res.* **88:** 2827-2837.

Downloaded from www.jlr.org by guest, on June 19, 2012

69. Stoffel, W., G. Blobel, and P. Walter. 1981. Synthesis in vitro and translocation of apolipoprotein A-I across microsomal vesicles. *Eur. J. Biochem.* **120:** 519–522.
70. Zannis, V. I., S. K. Karathanasis, H. T. Keutmann, G. Goldberger, and J. L. Breslow. 1983. Intracellular and extracellular processing of human apolipoprotein A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide. *Proc. Natl. Acad. Sci. USA.* **80:** 2574–2578.
71. Stoffel, W., C. Bode, and K. Knyrim. 1983. Serum apolipoprotein A-I synthesis in rat hepatocytes and its secretion as a proform. *Hoppe Seyler's Z. Physiol. Chem.* **364:** 439–445.
72. Banerjee, D., G. Grieninger, J. L. Parkes, T. M. Mukherjee, and C. M. Redman. 1986. Regulation of apoA-I processing in cultured hepatocytes. *J. Biol. Chem.* **261:** 9844–9849.
73. Banerjee, D., T. K. Mukherjee, and C. M. Redman. 1985. Biosynthesis of high density lipoprotein by chicken liver: intracellular transport and proteolytic processing of nascent apolipoprotein A-I. *J. Cell. Biol.* **101:** 1219–1226.
74. Yang, C-Y., Z. W. Gu, W. Patsch, S. Weng, T. W. Kim, and L. Chan. 1987. The complete amino acid sequence of proapolipoprotein A-I of chicken high density lipoproteins. *FEBS Lett.* **224:** 261–266.
75. Edelstein, C., J. I. Gordon, K. Toscas, H. F. Sims, A. W. Strauss, and A. M. Scanu. 1983. In vitro conversion of proapoprotein A-I to apoprotein A-I. Partial characterization of an extracellular enzyme activity. *J. Biol. Chem.* **258:** 11430–11433.
76. Gordon, J. I., H. F. Sims, C. Edelstein, A. M. Scanu, and A. W. Strauss. 1986. Extracellular processing of proapolipoprotein A-II in HepG2 cell cultures is mediated by a 54-kDa protease immunologically related to cathepsin B. *J. Biol. Chem.* **260:** 14824–14831.
77. Gordon, J. I., H. F. Sims, C. Edelstein, A. M. Scanu, and A. W. Strauss. 1984. Human proapolipoprotein A-II is cleaved following secretion from HepG2 cells by a thiol protease. *J. Biol. Chem.* **259:** 15556–15563.
78. Stoffel, W. 1984. Synthesis, transport, and processing of apolipoproteins of high density lipoproteins. *J. Lipid Res.* **25:** 1586–1592.
79. Gordon, J. I., H. F. Sims, A. W. Strauss, A. M. Scanu, C. Edelstein, and R. E. Byrne. 1986. Proteolytic processing and compartmentalization of the primary translation products of mammalian apolipoprotein mRNAs. *CRC Crit. Rev. Biochem.* **20:** 37–71.
80. Utermann, G., K. H. Vogelberg, A. Steinmetz, W. Schoenborn, N. Pruin, M. Jaeschke, M. Hees, and H. Canzler. 1979. Polymorphism of apolipoprotein E. II. Genetics of hyperlipoproteinemia type III. *Clin. Genet.* **15:** 37–62.
81. Zannis, V. I. 1986. Genetic polymorphism in human apolipoprotein E. *Method Enzymol.* **128:** 823–851.
82. Fojo, S. S., L. Taam, T. Fairwell, R. Ronan, C. Bishop, M. S. Meng, J. M. Hoeg, D. L. Sprecher, and H. B. Brewer, Jr. 1986. Human preproapolipoprotein C-II. Analysis of major plasma isoforms. *J. Biol. Chem.* **261:** 9591–9594.
83. Ghiselli, G., M. F. Rohde, S. Tanenbaum, S. Krishnan, and A. M. Gotto, Jr. 1985. Origin of apolipoprotein A-I polymorphism in plasma. *J. Biol. Chem.* **260:** 15662–15668.
84. Polacek, D., R. E. Byrne, G. M. Fless, and A. M. Scanu. 1986. In vitro proteolysis of human plasma low density lipoproteins by an elastase release from human blood polymorphonuclear cells. *J. Biol. Chem.* **261:** 2057–2063.
85. Hoeg, J. M., M. S. Meng, R. Ronan, T. Fairwell, and H. B. Brewer, Jr. 1986. Human apolipoprotein A-I, post-translational modification by fatty acid acylation. *J. Biol. Chem.* **261:** 3911–3914.
86. Davis, R. A., G. M. Clinton, R. A. Borchardt, M. Malone-McNeal, T. Tan, and G. R. Lattier. 1984. Intrahepatic assembly of very low density lipoproteins. Phosphorylation of small molecular weight apolipoprotein B. *J. Biol. Chem.* **259:** 3383–3386.
87. Reuben, M. A., and J. Elovson. 1985. Biosynthesis of apolipoprotein B in rat liver and intestine. *Fed. Proc.* **44:** 1455.
88. Marcel, Y. L., M. Hogue, R. Theolis, Jr., and R. W. Milne. 1982. Mapping of antigenic determinants of human apolipoprotein B using monoclonal antibodies against low density lipoproteins. *J. Biol. Chem.* **257:** 13165–13168.
89. Cardin, A. D., C. A. Price, N. Hirose, M. A. Krivanek, D. T. Blankenship, J. Chao, and S. J. T. Mao. 1986. Structural organization of apolipoprotein B-100 of human plasma low density lipoproteins. Comparison to B-48 of chylomicrons and very low density lipoproteins. *J. Biol. Chem.* **261:** 16744–16748.
90. Young, S. G., S. J. Bertico, T. M. Scott, B. W. Dubois, L. K. Curtiss, and J. L. Witztum. 1986. Parallel expression of the MB19 genetic polymorphism in apoprotein B-100 and apoprotein B-48. Evidence that both apoproteins are products of the same genes. *J. Biol. Chem.* **261:** 2995–2998.
91. Chan, L., P. VanTuinen, D. Ledbetter, S. P. Daiger, A. M. Gotto, Jr., and S. H. Chen. 1985. The human apolipoprotein B-100 gene: a highly polymorphic gene that maps to the short arm of chromosome 2. *Biochem. Biophys. Res. Commun.* **133:** 248–255.
92. Huang, L. S., D. A. Miller, G. A. P. Bruns, and J. L. Breslow. 1986. Mapping of the APOB gene to chromosome 2p and demonstration of a two-allele restriction fragment length polymorphism. *Proc. Natl. Acad. Sci. USA.* **83:** 644–648.
93. Mehrabian, M., R. S. Sparks, T. Mohandes, I. J. Klisak, V. N. Schumaker, C. Heinzman, S. Zollman, Y. Ma, and A. J. Lusis. 1986. Human apolipoprotein B: chromosomal mapping and DNA polymorphisms of hepatic and intestinal species. *Somatic Cell Mol. Genet.* **12:** 245–254.
94. Blackhart, B. D., E. M. Ludwig, V. R. Pierotti, L. Caiati, M. A. Onasch, S. C. Wallis, L. Pownell, R. Pease, T. J. Knott, M. L. Chu, R. W. Mahley, J. Scott, B. J. McCarthy, and B. Levy-Wilson. 1986. Structure of the human apolipoprotein B gene. *J. Biol. Chem.* **261:** 15364–15367.
95. Ludwig, E. H., B. D. Blackhart, V. R. Pierotti, L. Caiati, C. Fortier, T. Knott, J. Scott, R. W. Mahley, B. Levy-Wilson, and B. J. McCarthy. 1987. DNA sequence of the human apolipoprotein B gene. *DNA.* **6:** 363–372.
96. Powell, L. M., S. C. Wallis, R. J. Pease, Y. H. Edwards, T. J. Knott, and J. Scott. 1987. A novel form of tissue-specific RNA processing produces apolipoprotein B-48 in intestine. *Cell.* **50:** 831–840.
97. Chen, S-H., G. Habib, C-Y. Yang, Z-W. Gu, B. R. Lee, S-a. Weng, S. R. Silberman, S-J. Cai, J. P. Deslypere, M. Rosseneu, A. M. Gotto, Jr., W-H. Li, and L. Chan. 1987. Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. *Science.* **238:** 363–366.
98. Chan, L., A. Dugaiczyk, and A. R. Means. 1980. Molecular cloning of the gene sequences of a major apoprotein in avian very low density lipoproteins. *Biochemistry.* **19:** 5631–5637.
99. Wieringa, B., G. Ab, and M. Gruber. 1981. The nucleotide sequence of the very low density lipoprotein II mRNA from chicken. *Nucleic Acids Res.* **9:** 489–501.
100. Dugaiczyk, A., A. S. Inglis, P. M. Strike, R. W. Burley, W. G. Beattie, and L. Chan. 1981. Comparison of the nucleotide sequence of cloned DNA coding for an apolipo-

Downloaded from www.jlr.org by guest, on June 19, 2012

ASBMB JOURNAL OF LIPID RESEARCH

protein (apoVLDL-II) from avain blood and the amino acid sequence of an egg yolk protein (apovitellin I): equivalence of the two sequences. *Gene.* **14:** 175-182.

101. Wiskocil, R., P. Goldman, and R. G. Deeley. 1981. Cloning and structural characterization of an estrogen-dependent apolipoprotein gene. *J. Biol. Chem.* **256:** 9662-9667.
102. Meijlink, F. C. P. W., A. D. van het Schip, A. C. Arnberg, B. Wieringa, G. Ab, and M. Gruber. 1981. Structure of the chicken very low density lipoprotein II gene. *J. Biol. Chem.* **256:** 9668-9671.
103. van het Schip, A., F. C. R. W. Meijlink, R. Stryker, M. Gruber, A. J. van Vliet, J. A. M. van de Klundert, and G. Ab. 1983. The nucleotide sequence of the chicken apo very low density lipoprotein II gene. *Nucleic Acids Res.* **11:** 2529-2540.
104. Tsao, Y-K., C-F. Wei, D. L. Robberson, A. M. Gotto, Jr., and L. Chan. 1985. Isolation and characterization of the human apolipoprotein A-II gene. Electron microscopic analysis of RNA:DNA hybrids, nucleotide sequence, identification of a polymorphic MspI site, and general structural organization of apolipoprotein genes. *J. Biol. Chem.* **260:** 15222-15231.
105. Das, H. K., J. McPherson, G. A. P. Bruns, S. K. Karathanasis, and J. L. Breslow. 1985. Isolation, characterization and mapping to chromosome 19 of the human apolipoprotein E gene. *J. Biol. Chem.* **260:** 6240-6247.
106. Elshourbagy, N. A., D. W. Walker, M. S. Boguski, J. I. Gordon, and J. M. Taylor. 1986. The nucleotide and derived amino acid sequence of human apolipoprotein A-IV mRNA and the close linkage of its gene to the genes of apolipoproteins A-I and C-III. *J. Biol. Chem.* **261:** 1998-2002.
107. Lin-Lee, Y. C., F. T. Kao, P. Cheung, and L. Chan. 1985. Apolipoprotein E gene mapping and expression: localization of the structural gene to human chromosome 19 and expression of apoE mRNA in lipoprotein and non-lipoprotein-producing tissues. *Biochemistry.* **24:** 3751-3756.
108. Jackson, C. L., G. A. P. Bruns, and J. L. Breslow. 1984. Isolation and sequence of a human apolipoprotein C-II cDNA clone and its use to isolate and map to human chromosome 19 the gene for apolipoprotein C-II. *Proc. Natl. Acad. Sci. USA.* **81:** 2945-2949.
109. Karathanasis, S. K., I. Yunis, and V. I. Zannis. 1986. Structure, evolution, and tissue-specific synthesis of human apolipoprotein A-IV. *Biochemistry.* **25:** 3962-3970.
110. Paik, Y. K., D. J. Chang, C. A. Reardon, G. E. Davis, R. W. Mahley, and J. M. Taylor. 1985. Nucleotide sequence and structure of the human apolipoprotein E gene. *Proc. Natl. Acad. Sci. USA.* **82:** 3445-3449.
111. Shelley, C. S., C. R. Sharpe, F. E. Baralle, and C. C. Shoulders. 1985. Comparison of the human apolipoprotein genes. ApoA-II presents a unique functional intron-exon junction. *J. Mol. Biol.* **186:** 43-51.
112. Wei, C-F., Y-K. Tsao, D. L. Robberson, A. M. Gotto, Jr., K. Brown, and L. Chan. 1985. The structure of the human apolipoprotein C-II gene. Electron microscopic analysis of RNA:DNA hybrids, complete nucleotide sequence, and identification of 5' homologous sequences among apolipoprotein genes. *J. Biol. Chem.* **260:** 15211-15221.
113. Cladaras, C., M. Hadzopoulou-Cladaras, R. T. Nolte, D. Atkinson, and V. I. Zannis. 1986. The complete sequence and structural analysis of human apolipoprotein B-100: relationship between apoB-100 and apo-48 forms. *EMBO J.* **5:** 3495-3507.
114. Deeb, S. S., A. G. Motulsky, and J. J. Albers. 1985. A partial cDNA clone for human apolipoprotein B. *Proc. Natl. Acad. Sci. USA.* **82:** 4983-4986.
115. Huang, L. S., S. C. Bock, S. I. Feinstein, and J. L. Breslow. 1985. Human apolipoprotein B cDNA clone isolation and demonstration that liver apolipoprotein B mRNA is 22 kilobases in length. *Proc. Natl. Acad. Sci. USA.* **82:** 6825-6829.
116. Knott, T. J., S. C. Rall, Jr., T. L. Innerarity, S. F. Jacobson, M. S. Urdea, B. Levy-Wilson, L. M. Powell, R. J. Pease, R. Eddy, H. Nakai, M. Byers, L. M. Priestley, E. Robertson, L. B. Rall, C. Betsholtz, T. B. Shows, R. W. Mahley, and J. Scott. 1985. Human apolipoprotein B: structure of carboxyl-terminal domains, sites of gene expression, and chromosomal localization. *Science.* **230:** 37-43.
117. Knott, J., R. J. Pease, L. M. Powell, S. C. Wallis, S. C. Rall, Jr., T. L. Innerarity, B. Blackhart, W. H. Taylor, Y. L. Marcel, R. Milne, D. Johnson, M. Fuller, A. J. Lusis, B. J. McCarthy, R. W. Mahley, B. Levy-Wilson, and J. Scott. 1986. Complete protein sequence and identification of structural domains of human apolipoprotein B. *Nature.* **323:** 734-738.
118. Law, S. W., K. J. Lackner, A. V. Hospattanakar, J. M. Anchors, A. Y. Sakaguchi, S. L. Naylor, and H. B. Brewer, Jr. 1985. Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of the gene to chromosome 2. *Proc. Natl. Acad. Sci. USA.* **82:** 8340-8344.
119. Law, S. W., S. M. Grant, K. Higuchi, A. Hospattankar, K. Lackner, N. Lee, and H. B. Brewer. 1986. Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino acid sequence. *Proc. Natl. Acad. Sci. USA.* **83:** 8142-8146.
120. Mehrabian, M., V. N. Schumaker, G. C. Fareed, R. R. West, D. F. Johnson, T. Kirchgessner, H. Lin, X. Wang, Y. Ma, E. Mendiaz, and A. J. Lusis. 1985. Human apolipoprotein B: identification of cDNA clones and characterization of mRNA. *Nucleic Acids Res.* **13:** 6937-6953.
121. Wei, C-F., S. H. Chen, C-Y. Yang, Y. L. Marcel, R. W. Milne, W-H. Li, J. T. Sparrow, A. M. Gotto, Jr., and L. Chan. 1985. Molecular cloning and expression of partial cDNAs and deduced amino acid sequence of a carboxyl-terminal fragment of human apolipoprotein B-100. *Proc. Natl. Acad. Sci. USA.* **82:** 7265-7269.
122. Yang, C-Y., S. H. Chen, S. H. Gianturco, W. A. Bradley, J. T. Sparrow, M. Tanimura, W-H. Li, D. A. Sparrow, H. DeLoof, M. Rosseneu, F. S. Lee, Z. W. Gu, A. M. Gotto, Jr., and L. Chan. 1986. Sequence, structure, receptor binding domains and internal repeats of human apolipoprotein B-100. *Nature.* **323:** 738-742.
123. Carlsson, P., S. O. Olofsson, G. Bondjers, C. Darnfors, O. Wiklund, and G. Bjursell. 1985. Molecular cloning of human apolipoprotein B cDNA. *Nucleic Acids Res.* **13:** 8813-8826.
124. Drayna, D. T., J. W. McLean, K. L. Wion, J. M. Trent, H. A. Drabkin, and R. M. Lawn. 1987. Human apolipoprotein D gene: gene sequence, chromosome localization, and homology to the α_{2u}-globulin superfamily. *DNA.* **6:** 199-124.
125. Boguski, M. S., N. Elshourbagy, J. M. Taylor, and J. I. Gordon. 1985. Comparative analysis of repeated sequences in rat apolipoprotein A-I, A-IV, and E. *Proc. Natl. Acad. Sci. USA.* **82:** 992-996.
126. Lusis, A. J., R. West, M. Mehrabian, M. A. Reuben, R. C. LeBoeuf, J. S. Kaptein, D. F. Johnson, V. N. Schumaker, M. P. Yuhasz, M. C. Schotz, and J. Elovson. 1985. Cloning and expression of apolipoprotein B, the major protein of low and very low density lipoproteins. *Proc. Natl. Acad. Sci. USA.* **82:** 4597-4601.

Downloaded from www.jlr.org by guest, on June 19, 2012

127. Lusis, A. J., B. A. Taylor, R. W. Wangenstein, and R. C. LeBoeuf. 1983. Genetic control of lipid transport in mice. II. Genes controlling structure of high density lipoproteins. *J. Biol. Chem.* **258:** 5071–5078.
128. Poncin, J. E., J. A. Martial, and J. E. Gielen. 1984. Cloning and structure analysis of the rat apolipoprotein A-I cDNA. *Eur. J. Biochem.* **140:** 493–498.
129. Rajavashisth, T. B., J. S. Kaptein, K. L. Reue, and A. J. Lusis. 1985. Evolution of apolipoprotein E: mouse sequence and evidence for an 11 nucleotide ancestral unit. *Proc. Natl. Acad. Sci. USA.* **82:** 8085–8089.
130. Fukazawa, C., A. Matsumoto, and J. M. Taylor. 1986. Complete nucleotide sequence of the gene encoding the rat apolipoprotein E. *Nucleic Acids Res.* **14:** 9527–9528.
131. Kroon, P. A., J. A. DeMartino, G. M. Thompson, and Y. S. Chao. 1986. Molecular cloning of partial cDNAs for rabbit liver apolipoprotein B and the regulation of its mRNA levels by dietary cholesterol. *Proc. Natl. Acad. Sci. USA.* **83:** 5071–5075.
132. Williams, S. C., S. M. Bruckheimer, A. J. Lusis, R. C. LeBoeuf, and A. J. Kinniburgh. 1986. Mouse apoprotein A-IV gene: nucleotide sequence and induction by a high lipid diet. *Mol. Cell Biol.* **6:** 3807–3814.
133. Datta, S., W-H. Li, I. Ghosh, C-C. Luo, and L. Chan. 1987. Structure and expression of dog apolipoprotein C-II and C-III mRNAs: implications for the evolution and functional constraints of apolipoprotein structure. *J Biol. Chem.* **262:** 10588–10593.
134. Kao, F. T., and L. Chan. 1986. Genetic mapping of apolipoprotein genes in the human genome by somatic cell hybrids. *Methods Enzymol.* **128:** 851–863.
135. Harper, M. E., and L. Chan. 1986. Chromosomal fine mapping of apolipoprotein genes by in situ nucleic acid hybridization to metaphase chromosomes. *Methods Enzymol.* **128:** 863–876.
136. Karathanasis, S. K. 1985. Apolipoprotein multigene family and tandem organization of human apolipoprotein A-I, C-III, and A-IV genes. *Proc. Natl. Acad. Sci. USA.* **82:** 6374–6378.
137. Cheung, P., F. T. Kao, M. L. Law, C. Jones, T. T. Puck, and L. Chan. 1984. Localization of the structural gene for human apolipoprotein A-I on the long arm of human chromosome 11. *Proc. Natl. Acad. Sci. USA.* **81:** 508–511.
138. Law, S. W., G. Gray, H. B. Brewer, Jr., A. Y. Sakaguchi, and S. L. Naylor. 1984. Human apolipoprotein A-I and C-III genes reside in the p11 → q13 region of chromosome 11. *Biochem. Biophys. Res. Commun.* **118:** 934–942.
139. Bruns, G. A. P., S. K. Karathanasis, and J. L. Breslow. 1984. Human apolipoprotein A-I/C-III gene complex is located on chromosome 11. *Arteriosclerosis.* **4:** 97–102.
140. Lusis, A. J., C. Heinzmann, R. S. Sparkes, J. Scott, T. J. Knott, R. Geller, M. C. Sparkes, and T. Mohandas. 1986. Regional mapping of human chromosome 19: organization of genes for plasma lipid transport (APOC1, -C2, and -E and LDLR) and the genes C3, PEPD, and GPI. *Proc. Natl. Acad. Sci. USA.* **83:** 3929–3933.
141. Knott, T. J., R. L. Eddy, M. E. Robertson, L. M. Priestley, J. Scott, and T. B. Shows. 1984. Chromosomal localization of the human apoprotein C-I gene and of a polymorphic apoprotein A-II gene. *Biochem. Biophys. Res. Commun.* **125:** 299–306.
142. Jeanpierre, M., D. Weil, M. C. Hors-Caylor, R. Williamson, C. Junien, and S. E. Humphries. 1984. The gene for apolipoprotein C-II is on human chromosome 19. *Somatic Cell Mol. Genet.* **10:** 645–649.
143. Myklebost, O., S. Rogne, B. Olaisen, T. Gedde-Dahl, Jr., and H. Prydz. 1984. The locus for apolipoprotein C-II is closely linked to the apolipoprotein E locus on chromosome 19 in man. *Hum. Genet.* **67:** 309–312.
144. Humphries, S. E., K. Berg, L. Gill, A. M. Cumming, F. W. Robertson, A. F. H. Stalenhoef, R. Williamson, and A. L. Barresen. 1984. The gene for apolipoprotein C-II is closely linked to the gene for apolipoprotein E on chromosome 19. *Clin. Genet.* **26:** 389–396.
145. Olaisen, B., P. Teisberg, and T. Gedde-Dahl, Jr. 1982. The locus for apolipoprotein E (apoE) is linked to the complement component 3 (C3) locus on chromosome 19 in man. *Hum. Genet.* **62:** 233–236.
146. Berg, K., and A. Heiberg. 1978. Linkage between familial hypercholesterolemia with xanthomatosis and the C3 polymorphism confirmed. *Cytogenet. Cell Genet.* **22:** 621–623.
147. Francke, U., M. S. Brown, and J. L. Goldstein. 1984. Assignment of the human gene for the low density lipoprotein receptor to chromosome 19: synteny of a receptor, a ligand, and a genetic disease. *Proc. Natl. Acad. Sci. USA.* **81:** 2826–2830.
148. Lackner, K. W., S. W. Law, H. B. Brewer, Jr., A. Y. Sakaguchi, and S. L. Naylor. 1984. The human apolipoprotein A-II gene is located on chromosome 1. *Biochem. Biophys. Res. Commun.* **122:** 877–883.
149. Tate, V., M. Finer, H. Boedtker, and P. Doty. 1982. Procollagen genes: further sequence studies and interspecies comparisons. *Cold Spring Harbor Symp. Quant. Biol.* **49:** 1039–1049.
150. Alexander, F., P. R. Young, and S. M. Tilghman. 1984. Evolution of the albumin: alpha-fetoprotein ancestral gene from the amplification of a 27 nucleotide sequence. *J. Mol. Biol.* **173:** 159–176.
151. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature.* **302:** 575–581.
152. Stein, J. P., J. Catterall, P. Kristo, A. R. Means, and B. W. O'Malley. 1980. Ovomucoid intervening sequences specify functional domains and generate protein polymorphism. *Cell.* **21:** 681–687.
153. Nathans, J., and D. S. Hogness. 1983. Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. *Cell.* **34:** 807–814.
154. Liscum, L., J. Finer-Moore, R. M. Stroud, K. L. Luskey, M. S. Brown, and J. L. Goldstein. 1985. Domain structure of 3-hydroxy-3-methylglutaryl coenzyme A reductase, a glycoprotein of the endoplasmic reticulum. *J. Biol. Chem.* **260:** 522–530.
155. Lonberg, N., and W. Gilbert. 1985. Intron/exon structure of the chicken pyruvate kinase gene. *Cell.* **40:** 81–90.
156. Gilbert, W. 1978. Why genes in pieces? *Nature.* **271:** 501.
157. Craig, C. S., S. Sprang, R. Sletterick, and W. J. Rutter. 1982. Intron-exon splice junctions map at protein surfaces. *Nature.* **299:** 180–182.
158. Chou, P. Y., and G. D. Fasman. 1978. Empirical predictions of protein conformations. *Annu. Rev. Biochem.* **47:** 251–276.
159. Eisenberg, D., R. M. Weiss, and T. C. Terwilliger. 1982. The helical hydrophobic moment: a measure of the amphiphilicity of a helix. *Nature.* **299:** 371–374.
160. McLachlan, A. D. 1977. Repeated helical pattern in apolipoprotein A-I. *Nature.* **267:** 465–466.
161. Lux, S. E., R. Hirz, R. I. Shrager, and A. M. Gotto, Jr. 1972. The influence of lipid on the conformation of human plasma high density apolipoproteins. *J. Biol. Chem.* **247:** 2598–2606.
162. Jackson, R. L., J. D. Morrisett, A. M. Gotto, Jr., and J. P.

Downloaded from www.jlr.org by guest, on June 19, 2012

ASBMB JOURNAL OF LIPID RESEARCH

Segrest. 1975. The mechanism of lipid-binding by plasma lipoproteins. *Mol. Cell. Biochem.* **6:** 43–50.
163. Jackson, R. L., J. D. Morrisett, J. T. Sparrow, J. P. Segrest, H. J. Pownall, L. C. Smith, H. F. Hoff, and A. M. Gotto, Jr. 1974. The interaction of apolipoprotein-serine with phosphatidylcholine. *J. Biol. Chem.* **249:** 5314–5320.
164. Segrest, J. P., R. L. Jackson, J. D. Morrisett, and A. M. Gotto, Jr. 1974. A molecular theory of lipid-protein interactions in the plasma lipoproteins. *FEBS Lett.* **38:** 247–258.
165. Doolittle, R. F. 1981. Similar amino acid sequences: chance or common ancestry? *Science.* **214:** 149–159.
166. Gotto, A. M., Jr., R. I. Levy, and D. S. Frederickson. 1968. Observations on the conformation of human beta-lipoprotein: evidence of the occurrence of beta structure. *Proc. Natl. Acad. Sci. USA.* **60:** 1436–1441.
167. Chen, G. C., and J. P. Kane. 1975. Temperature dependence of the optical activity of human serum low density lipoprotein. The role of lipids. *Biochemistry.* **14:** 3357–3362.
168. Osterman, D., R. Mora, F. J. Kezdy, E. T. Kaiser, and S. C. Meredit. 1984. A synthetic amphiphilic beta-strand tridecapeptide: a model for apolipoprotein B. *J. Am. Chem. Soc.* **106:** 6845–6847.
169. Yang, C-Y., T. W. Kim, Q. Pao, L. Chan, R. D. Knapp, A. M. Gotto, Jr., and H. J. Pownall. 1988. Structure and conformational analysis of lipid-associating peptides of apolipoprotein B-100 produced by trypsinolysis. Submitted for publication.
170. Fukushima, D., J. P. Kupferberg, S. Yokoyama, D. J. Kroon, E. T. Kaiser, and F. J. Kezdy. 1979. A synthetic amphipathic helical docasapeptide with the surface properties of plasma apolipoprotein A-I. *J. Am. Chem. Soc.* **101:** 3703–3704.
171. Kaiser, E. T., and F. J. Kezdy. 1983. Secondary structures of proteins and peptides in amphipathic environments (a review). *Proc. Natl. Acad. Sci. USA.* **80:** 1137–1143.
172. Nakagawa, S. H., H. S. H. Lau, F. J. Kezdy, and E. T. Kaiser. 1985. The use of polymer-bound oximes for the synthesis of large peptides usable in segment condensation: synthesis of a 44 amino acid amphipathic peptide model of apolipoprotein A-I. *J. Am. Chem. Soc.* **107:** 7087–7092.
173. Sparrow, J. T., and A. M. Gotto, Jr. 1982. Apolipoprotein/lipid interactions: studies with synthetic polypeptides. *CRC Crit. Rev. Biochem.* **13:** 87–107.
174. Rajavashisth, T. B., P. A. Dawson, D. L. Williams, J. E. Shackelford, H. Lebherz, and A. J. Lusis. 1987. Structure, evolution and regulation of chicken apolipoprotein A-I. *J. Biol. Chem.* **262:** 7058–7065.
175. Byrnes, L., C-C. Luo, W-H. Li, C-Y. Yang, and L. Chan. 1987. Chicken apolipoprotein A-I: cDNA sequence, expression and evolution. *Biochem. Biophys. Res. Commun.* **148:** 485–492.
176. Pontes, M., X. Xu, D. Graham, M. Riley, and R. F. Doolittle. 1987. cDNA sequences of two apolipoproteins from lamprey. *Biochemistry.* **26:** 1611–1617.
177. Blumberg, B. S., S. Dray, and J. C. Robinson. 1962. Antigen polymorphism of a low density beta-lipoprotein allotype in human serum. *Nature.* **194:** 656–658.
178. Butler, R., E. Butler, R. Scherz, and R. Pflugshaupt. 1978. *In* Proceedings of the 25th Colloquium on Protides of the Biological Fluids. H. Peeters, editor. Pergamon Press, New York. 255–262.
179. Berg, K. 1979. *In* The Biochemistry of Atherosclerosis: Inherited Lipoprotein Variation and Atherosclerotic Disease. A. M. Scanu, R. W. Wissler, and G. S. Getz, editors. Dekker, New York. 419–480.
180. Schumaker, V. N., M. T. Robinson, L. K. Curtiss, R. Butler, and R. S. Sparkes. 1984. Anti-apoprotein B monoclonal antibodies detect human low density lipoprotein polymorphism. *J. Biol. Chem.* **259:** 6423–6430.
181. Rapacz, J., J. Haster-Rapacz, K. M. Taylor, W. J. Checovich, and A. D. Attie. 1986. Lipoprotein mutations in pigs are associated with elevated plasma cholesterol and atherosclerosis. *Science.* **234:** 1573–1577.
182. Margolis, S., and R. G. Langdon. 1966. Studies on human serum, beta$_1$-lipoprotein. III. Enzymatic modifications. *J. Biol. Chem.* **241:** 485–493.
183. Chapman, M. J., S. Goldstein, and G. L. Mills. 1978. Limited tryptic digestion of human serum low density lipoprotein. Isolation and characterization of the protein-deficient particles and of its apoproteins. *Eur. J. Biochem.* **87:** 475–488.
184. DeLoof, H., M. Rosseneu, C-Y. Yang, W-H. Li, A. M. Gotto, Jr., and L. Chan. 1987. Human apolipoprotein B: analysis of internal repeats and homology with other apolipoproteins. *J. Lipid Res.* **28:** 1455–1465.
185. Fielding, P. E., and C. J. Fielding. 1980. A cholesteryl ester transfer complex in human plasma. *Proc. Natl. Acad. Sci. USA.* **77:** 3327–3330.
186. Albers, J. J., M. C. Cheung, S. L. Ewens, and J. H. Tollefson. 1981. Characterization and immunoassay of apolipoprotein D. *Atherosclerosis.* **39:** 395–409.
187. Utermann, G., H. J. Menzel, G. Adler, P. Decker, and W. Weber. 1980. Substitution in vitro of lecithin:cholesterol acyltransferase. Analysis of changes in plasma lipoproteins. *Eur. J. Biochem.* **107:** 225–241.
188. Li, W-H., C-I. Wu, and C-C. Luo. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2:** 150–174.
189. Dayhoff, M. O. 1972. Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Silver Spring, MD.
190. Kimura, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge, England.
191. Li, W-H., C-C. Luo, and C-I. Wu. 1985. Evolution of DNA sequences. *In* Molecular Evolutionary Genetics. R. J. MacIntyre, editor. Plenum Press, New York. 1–94.
192. Romer, A. S. 1966. Vertebrate Paleontology. University of Chicago Press, Chicago.
193. Fitch, W. M., and E. Margoliash. 1967. Construction of phylogenetic trees. A method based on mutation distances as estimated from cytochrome C sequences is of general applicability. *Science.* **155:** 279–284.
194. Wu, C-I., and W-H. Li. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. USA.* **82:** 1741–1745.
195. Boguski, M. S., E. H. Berkenmeier, N. A. Elshourbagy, J. M. Taylor, and J. I. Gordon. 1986. Evolution of the apolipoproteins. Structure of the rat apoA-IV gene and its relationship to the human genes for apoA-I, C-III and E. *J. Biol. Chem.* **261:** 6398–6407.
196. Fitch, W. M., T. Smith, and J. L. Breslow. 1986. Detecting internally repeated sequences and inferring the history of duplication. *Methods Enzymol.* **128:** 773–788.
197. Chapman, M. J., S. Goldstein, G. L. Mills, and C. Leger. 1978. Distribution and characterization of the serum lipoproteins and their apoproteins in the rainbow trout (*Salmo gairdnerii*). *Biochemistry.* **17:** 4455–4464.
198. Chapman, M. J. 1980. Animal lipoproteins: chemistry, structure, and comparative aspects. *J. Lipid Res.* **21:** 789–853.
199. Babin, P. J. 1987. Apolipoproteins and the association of egg yolk proteins with plasma high density lipoproteins af-

ASBMB JOURNAL OF LIPID RESEARCH JLR
Downloaded from www.jlr.org by guest, on June 19, 2012

ter ovulation and follicular atresia in the rainbow trout (*Salmo gairdnerii*). *J. Biol. Chem.* **262:** 4290–4296.
200. Zorich, N., A. Jonas, and H. L. Pownall. 1985. Activation of lecithin:cholesterol acyltransferase by human apolipoprotein E in discoidal complex with lipids. *J. Biol. Chem.* **260:** 8831–8837.
201. Steinmetz, A., H. Kaffarnik, and G. Utermann. 1985. Activation of phosphatidylcholine-sterol acyltransferase by human apolipoprotein E isoforms. *Eur. J. Biochem.* **152:** 747–751.
202. Rall, S. C., Jr., K. H. Weisgraber, and R. W. Mahley. 1982. Human apolipoprotein E. The complete amino acid sequence. *J. Biol. Chem.* **257:** 4171–4178.
203. Connelly, P. W., G. F. T. Hoffman, and J. A. Little. 1987. Structure of apolipoprotein C-$II_{Toronto}$: a non-functional human apolipoprotein. *Proc. Natl. Acad. Sci. USA.* **84:** 270–273.
204. Kinnunen, P. K. J., R. L. Jackson, L. C. Smith, A. M. Gotto, Jr., and J. T. Sparrow. 1977. Activation of lipoprotein lipase by native and synthetic fragments of human plasma apoC-II. *Proc. Natl. Acad. Sci. USA.* **74:** 4848–4852.
205. Smith, L. C., P. K. J. Kinnunen, R. L. Jackson, A. M. Gotto, Jr., and J. T. Sparrow. 1978. Activation of lipoprotein lipase by native and synthetic peptide fragments of apolipoprotein C-II. *In* Proceedings of International Conference on Atherosclerosis. L. A. Carlson, R. Paoletti, and G. Weber, editors. Raven Press, New York. 269–273.
206. Edelstein, C., L. L. Lewis, J. R. Shainoff, H. Naito, and A. M. Scanu. 1976. Isolation and characterization of a dog serum lipoprotein having apolipoprotein A-I as its predominant protein constituent. *Biochemistry.* **15:** 1934–1941.
207. Jackson, R. L., H. N. Baker, O. D. Taunton, L. C. Smith, C. W. Garner, and A. M. Gotto, Jr. 1973. A comparison of the major apolipoprotein from pig and human high density lipoproteins. *J. Biol. Chem.* **248:** 2639–2644.
208. Jonas, A. 1972. Physicochemical properties of bovine serum high density lipoprotein. *J. Biol. Chem.* **247:** 7767–7772.
209. Hillyard, L. A., H. M. White, and S. A. Pangburn. 1972. Characterization of apolipoproteins in chicken serum and egg yolk. *Biochemistry.* **11:** 511–518.
210. Sparrow, J. T., and A. M. Gotto, Jr. 1980. Phospholipid binding studies with synthetic apolipoprotein fragments. *Ann. N.Y. Acad. Sci.* **348:** 187–211.
211. Pownall, H. J., A. Hu, A. M. Gotto, Jr., J. J. Albers, and J. T. Sparrow. 1980. Activation of lecithin:cholesterol acyltransferase by a synthetic model lipid-associating peptide. *Proc. Natl. Acad. Sci. USA.* **77:** 3154–3158.
212. Yokoyama, S., D. Fukushima, J. P. Kupferberg, F. J. Kezdy, and E. T. Kaiser. 1980. The mechanism of activation of lecithin:cholesterol acyltransferase by apolipoprotein A-I and an amphipathic peptide. *J. Biol. Chem.* **255:** 7333–7339.
213. Mahley, R. W., and T. L. Innerarity. 1983. Lipoprotein receptors and cholesterol homeostasis. *Biochim. Biophys. Acta.* **737:** 197–222.
214. Innerarity, T. L., E. J. Friedlander, S. C. Rall, Jr., K. H. Weisgraber, and R. W. Mahley. 1983. The receptor binding domain of human apolipoprotein E: binding of apolipoprotein E fragments. *J. Biol. Chem.* **258:** 12341–12347.
215. Weisgraber, K. H., T. L. Innerarity, K. J. Harder, R. W. Mahley, R. W. Milne, Y. L. Marcel, and J. T. Sparrow. 1983. The receptor binding domain of human apolipoprotein E. Monoclonal antibody inhibition of binding. *J. Biol. Chem.* **258:** 12348–12354.
216. Weisgraber, K. H., T. L. Innerarity, and R. W. Mahley. 1982. Abnormal lipoprotein receptor binding activity of the human apoprotein due to cysteine–arginine interchange at a single site. *J. Biol. Chem.* **257:** 2518–2521.
217. Weisgraber, K. H., S. C. Rall, Jr., T. L. Innerarity, R. W. Mahley, T. Kuusi, and C. Ehnholm. 1984. A novel electrophoretic variant of human apolipoprotein E: identification and characterization of apolipoprotein E1. *J. Clin. Invest.* **73:** 1024–1033.
218. Rall, S. C., Jr., K. H. Weisgraber, T. L. Innerarity, and R. W. Mahley. 1982. Structural basis for receptor binding heterogeneity of apolipoprotein E from the type III hyperlipoproteinemic subjects. *Proc. Natl. Acad. Sci. USA.* **79:** 4696–4700.
219. Innerarity, T. L., K. H. Weisgraber, K. S. Arnold, S. C. Rall, Jr., and R. W. Mahley. 1984. Normalizaiton of receptor binding of apolipoprotein E2. *J. Biol. Chem.* **259:** 7261–7267.
220. Brunzell, J. D., A. D. Sniderman, J. J. Albers, and P. O. Kwiterovich, Jr. 1984. Apoproteins B and A-I and coronary artery disease in humans. *Atherosclerosis.* **4:** 79–83.
221. Sniderman, A., S. Shapiro, D. Marpole, B. Skinner, B. Teng, and P. O. Kwiterovich, Jr. 1980. Association of coronary atherosclerosis with hyperbetalipoproteinemia (increased protein but normal cholesterol levels in human plasma low density lipoproteins). *Proc. Natl. Acad. Sci. USA.* **77:** 604–608.
222. Matsumoto, A., H. Aburatani, Y. Shibasaki, T. Kodama, F. Takake, and H. Itakura. 1987. Cloning and regulation of rat apolipoprotein B mRNA. *Biochem. Biophys. Res. Commun.* **142:** 92–99.
223. Goldstein, S., M. J. Chapman, and G. L. Mills. 1977. Biochemical and immunological evidence for the presence of an apolipoprotein B-like component in the serum low density lipoproteins of several animal species. *Atherosclerosis.* **28:** 93–100.
224. Schiffer, M., and A. B. Edmundson. 1967. Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. *Biophys. J.* **7:** 121–125.
225. Staden, R. 1982. An iterative graphics program for comparing and aligning nucleic acid and amino acid sequences. *Nucleic Acids Res.* **10:** 2951–2961.
226. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157:** 105–132.

ASBMB JOURNAL OF LIPID RESEARCH JLR
Downloaded from www.jlr.org by guest, on June 19, 2012