Structure and expression of dog apolipoprotein A-I, E, and C-I mRNAs: implications for the evolution and functional constraints of apolipoprotein structure

Chi-Cheng Luo,* Wen-Hsiung Li,* and Lawrence Chan†

The 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

Abstract Dog apolipoprotein (apo) C-I, A-I, and E cDNA clones were identified in a dog liver cDNA library in λgt10 by hybridization to synthetic oligonucleotide probes with the corresponding human DNA sequences. The longest clone for each apolipoprotein was completely sequenced. The apoC-I cDNA sequence predicts a protein of 62 residue mature peptide preceded by a 26 amino acid signal peptide. The apoA-I cDNA sequence predicts a 242 residue mature peptide, a 6 residue pro-segment, and an 18 residue signal peptide. The apoE cDNA, which lacks the signal peptide region, predicts a mature peptide of 291 amino acid residues. Slot blot hybridization of total RNA isolated from various dog tissues to dog apoC-I, A-I, and E cDNA probes indicates that apoC-I mRNA is detectable in liver only, apoA-I mRNA is present in liver and small intestine, though the concentration in the latter tissue is only ~15% of that in the liver, and apoE mRNA is present in multiple tissues including liver, jejunum, urinary bladder, ileum, colon, brain, kidney, spleen, pancreas, and testis with relative concentrations (%) of 100, 17.5, 7.5, 6.9, 5.9, 5.5, 5.0, 3.3, 1.0, and 1.0, respectively. These tissue distributions indicate that nascent lipoprotein particles produced in the dog small intestine would contain apoA-I and apoE but not apoC-I. The widespread tissue distribution of apoE mRNA indicates that like other mammals, peripheral synthesis of apoE contributes significantly to the total apoE pool in dog. We next compared the cDNA sequences among different vertebrate species for apoC-I (human and dog), A-I (human, rat, dog, rabbit and chicken), and E (human, rat, dog and rabbit) and calculated the rate of nucleotide substitution for each gene. Our results indicate that apoC-I has evolved rather rapidly and that on the whole, apoA-I is more conservative than apoE, contradictory to an earlier suggestion. ApoA-I is also more conservative than a region (residues 4204–4536) at the carboxyl-terminal portion, but less conservative than a region (residues 595–979) at the amino-terminal portion of apoB-100. Some regions in each of the apolipoproteins studied are better conserved than others and the rate of evolution of individual regions seems to be related to the stringency of functional requirements. Finally, we estimate that the human apoC-I pseudogene arose more than 35 million years ago, becoming nonfunctional soon after its formation. —**Luo, C-C., W-H. Li, and L. Chan.** Structure and expression of dog apolipoprotein A-I, E, and C-I mRNAs: implications for the evolution and functional constraints of apolipoprotein structure. *J. Lipid Res.* 1989. **30:** 1735–1746.

Supplementary key words cDNA clones • tissue distribution • apoC-I pseudogene

The plasma lipoproteins are macromolecular complexes of lipids (triacylglycerols, cholesterol, and phospholipids) and protein. They are the vehicles for the transport of the hydrophobic lipid moieties from one tissue to another for metabolism. The protein components of plasma lipoproteins are known as apolipoproteins. All apolipoproteins share the ability to spontaneously bind lipid. In addition, many of them have acquired highly specialized functions (for review, see refs. 1, 2).

In this communication, we examined the structure and expression of the mRNA for three apolipoproteins in the dog, an animal used extensively as a model for cardiovascular and lipoprotein research (e.g., 3–5). The three canine apolipoproteins we examined here are apolipoprotein (apo) A-I, apoE, and apoC-I. ApoA-I is the major protein in high density lipoproteins, whose concentration is inversely related to the propensity for development of atherosclerosis (6–9). It is the major activator for the enzyme lecithin:cholesterol acyltransferase (LCAT) (10–12). ApoE is a constituent of chylomicrons, chylomicron remnants, very low density lipoproteins, and special classes of high density lipoproteins with apoE (HDL_1, HDL_c). It is an interesting protein in that it confers many unique functions to the lipoprotein particle, e.g., high affinity binding to the LDL receptor and a specific apoE receptor (13–15). The protein is synthesized in numerous tissues (16–20) and may be involved in such diverse functions as reverse cholesterol transport (16), neuronal regeneration (20), and immunomodulation (21) (for review, see ref. 22). ApoC-I is the smallest of the apolipoproteins. It also has the ability to activate LCAT in vitro (23).

Abbreviations: apo, apolipoprotein(s); LCAT, lecithin:cholesterol acyltransferase; HDL, high density lipoproteins.

While the structures of apoA-I and apoE have been determined in multiple species (human, rat, rabbit, and chicken for apoA-I, and human, rat, and rabbit for apoE) (see ref. 2 for references therein), the structure of apoC-I is known for humans only. The availability of the apolipoprotein mRNA sequences for apoA-I, C-I, and E from another species has allowed us to closely examine the evolution of these interesting proteins, and to infer the structure–function relationship in each of them. It also enables us to estimate the time of appearance of the human apoC-I pseudogene.

Although a number of laboratories have studied lipoprotein metabolism in the dog (3–5), the contribution of various tissues to the total apolipoprotein production in this animal is unknown. In the present communication, we present the distribution of apoA-I, E, and C-I mRNA among different dog tissues. The information is important to our understanding of lipoprotein metabolism in the canine model.

MATERIALS AND METHODS

Restriction enzymes were from BRL (Bethesda Research Laboratories), IBI (International Biotechnologies, Inc.), and BM (Boehringer Mannheim). DNA polymerase I and the Klenow fragment of polymerase I were from BM. Avian myeblastosis virus DNA polymerase (reverse transcriptase), T4 DNA polymerase, and T4 DNA ligase were from BRL. Proteinase K was from Merck. Deoxyribonucleotides and dideoxyribonucleotides were from Amersham. The ^{32}P- and ^{35}S-labeled mononucleotide triphosphates were from ICN or Amersham.

Total and polyA RNA isolation from dog tissues

Total RNA was prepared from various dog tissues, immediately after killing by the guanidinium thiocyanate method (24). Total RNA was purified from the intial crude RNA extract by cesium chloride step gradient (24). Quality of the RNA was checked by agarose gel electrophoresis. The total RNA was passed over an oligo-dT cellulose column twice to obtain polyA mRNA (25).

Construction of cDNA library, identification and sequence analysis of cDNA clones

A dog hepatic cDNA library was constructed in the phage vector λgt10 by the method of Gubler and Hoffman (26) using oligodeoxythymidylate primers and EcoRI linker ligation for insertion of the cDNA into the EcoRI site of λgt10. The cDNA library was screened by plaque hybridization by using standard procedures (27). Oligonucleotide probes (two 21-mers, with sequences TTCTGGCAGCAAGATGAACCC and GAGAAGGCCAAACCCGCGCTC, corresponding to nucleotide positions 61–81 and 685–705, for human apoA-I; two 21-mers, with sequences CGCTTTTGGGATTACCTGCGC and GTGGAAGACATGCAGCGCCAG, corresponding to nucleotide positions 148–168 and 859–879 for human apoE; and one 37-mer with the sequence TGAAATGTCTCTGAAAACCACTCCCGCATCTTGGCAG corresponding to nucleotide positions 221–185 in the antisense strand for human apoC-I) were synthesized on an Applied Biosystems Model 380A DNA synthesizer. These were used to identify the corresponding dog cDNA clones by cross-hybridization. The temperature of hybridization was 50°C. For apoA-I and apoE, the primary screening was each performed with two 21-base oligonucleotides. During secondary screening, duplicate filters were screened with individual oligonucleotides. Plaques were purified by secondary and tertiary screening. cDNA inserts were recovered from the clones by digestion with EcoRI. They were subcloned in both pGEM-Blue and M13 vectors.

The nucleotide sequences of the cloned dog apoA-I, apoC-I, and apoE cDNAs were determined by the dideoxy nucleotide chain termination method (28). The cDNA inserts were subcloned into the EcoRI sites of the M13 phage vectors mp18 or mp19 before sequencing. Sequencing was carried out on both strands using the M13 universal primer. Synthetic oligonucleotide primers were used for sequencing internal regions of each clone.

Northern blot and slot blot analysis of apoA-I, apoC-I and apoE mRNAs

For Northern blot analysis, 20 μg of total RNA was denatured by heating at 70°C in 50% formaldehyde, then subjected to electrophoresis for 3 h at 70 volts on a 1% agarose gel in 6% formaldehyde, 50 mM HEPES, pH 7.8, 1 mM EDTA. After electrophoresis, the gels were rinsed twice in water for 15 min each, then washed in 2 × SSC. The RNA was transferred to a Gene Screen membrane (NEN, Du Pont Company) in 10 × SSC overnight. For slot blot analysis, varying amounts (1–20 μg) of total RNA from various canine organs were directly blotted onto nitrocellulose paper using a slot blot apparatus (Schleicher and Schuell). The double-stranded cloned canine apoA-I, E, and C-I cDNA inserts (purified from their respective pGEM-blue vectors) were labeled with [^{32}P]dNTP by nick-translation or random oligonucleotide priming. Prehybridization, hybridization to ^{32}P-labeled nick-translated probes, and washing were as described (19). The Northern blots and slot blots were exposed to Kodak X-ray film, XAR-5, for 18–30 h. Autoradiograms were scanned with a MacBeth TD932 densitometer. Relative concentrations of the respective mRNAs were calculated from the linear regression coefficients (slopes) deduced from the signals obtained with different RNA concentrations applied to the blot.

Statistical analysis of nucleotide substitution rates

In estimating the number of nucleotide substitutions between two genes, we have used the method of Li, Wu, and Luo (29). In this method, nucleotide sites and substitutions are classified as synonymous (causing no amino acid change) and nonsynonymous. For example, the first two positions of the codon UUU are nonsynonymous, while the third position is counted as one-third synonymous and two-thirds nonsynonymous. The method gives the number (K_S) of (synonymous) substitutions per synonymous site and the number (K_A) of (nonsynonymous) substitutions per nonsynonymous site.

RESULTS

cDNA cloning and deduced amino acid sequence of dog apoA-I, apoC-I, and apoE

Using oligonucleotide probes from the corresponding human cDNA sequences for cross hybridization, we identified 23, 14, and 12 clones of dog apoA-I, apoC-I, and apoE cDNAs, respectively, in a dog liver library in λgt10. The longest cDNA clones were completely sequenced. A partial restriction map of the canine cDNA clones and the sequencing strategy are shown in **Fig. 1**.

ApoA-I. An 883 base pair apoA-I clone (λAI-11) was isolated from the dog liver library and its nucleotide sequence was determined (**Fig. 2**). Analysis of this sequence revealed an open reading frame of 798 nucleotides, flanked by 5'- and 3'-untranslated regions of 12 and 73 nucleotides, respectively. Fourteen bases upstream of the polyA tail is a putative polyadenylation signal, AACAAA, a variant of the canonical signal sequence, AATAAA.

The first 24 residues of the derived amino acid sequence contain the signal peptide (18 residues) and prosegment (6 residues). They show a high degree of homology to the corresponding human, chicken, rat, and rabbit sequences. Like both the human and rat apoA-I, the prosegment in dog also contains a Gln-Gln dipeptide, unusual amino acids for protein precursors that are processed proteolytically (30, 31). It is interesting that the rabbit prosequence ends with Gln-Arg, and the chicken with Gln-His. This indicates that the requirement for Gln-Gln next to cleavage site is not absolute. The mature peptide of dog apoA-I contains 242 residues. It is one residue shorter than the corresponding human sequence. Codon 3 that encodes proline is duplicated in human apoA-I. Only a single proline residue is present in this position in all other known vertebrate apoA-Is (see below). The predicted amino acid sequence of canine apoA-I differs from a previously reported sequence determined on the purified protein in that amino acid 211 is Glu instead of Gln in the latter sequence (32).

ApoC-I. The dog apoC-I clone (λCI-7) is 427 bp in length plus a polyA tail. The DNA sequence includes 23 bp in the 5'-untranslated region, 264 bp in the coding region, a termination codon (TGA), and a 3'-untranslated region of 137 bp (**Fig. 3**). A polyadenylation signal sequence AATAAA precedes the polyA tail by 12 bases.

The DNA-deduced amino acid sequence contains 62 residues of dog apoC-I mature peptide region, preceded by a 26 amino acid signal peptide. Dog apoC-I is thus longer than the corresponding human protein by 5 amino acid residues. The additional residues in the dog sequence occur at positions 9–12 (4 residues) and at position 61 (see below).

ApoE. On an initial screening, the longest canine apoE cDNA clone that was identified in the library was 353 bp in length. It spans the 3' region of the molecule including all sequences 5' upstream to the codon that encodes amino acid Ala-174. On subsequent screening, we identified several additional apoE cDNA clones, the longest of which, designated λE-12, encodes the entire mature peptide region but still misses the signal peptide region. The nucleotide and deduced amino acid sequence of this clone is shown in **Fig. 4**. λE-12 contains 1,019 nucleotides. It includes an open reading frame of 873 bp and a 3'-untranslated region of 146 bp plus the polyA tail. It predicts a mature polypeptide of 291 amino acid residues, 8 residues shorter than the human sequence. Residues 121–150 of the sequence show a complete match to a partial sequence of purified canine apoE reported by Weisgraber et al. (33).

ApoC-I, A-I and E mRNA expression in various dog tissues

The dog has been used as an experimental animal for lipoprotein metabolism. However, little is known concerning the site of synthesis of some of these proteins. The isolation of cDNAs of dog apoC-I, A-I, and E mRNA allowed us to quantify the individual mRNAs from different tissues. Analysis of total RNAs extracted from the liver, small intestine, pancreas, brain, lung, spleen, kidney, urinary bladder, and testis blotted on nitrocellulose paper and hybridized to the respective ^{32}P-labeled cDNAs revealed that apoC-I mRNA is detected only in the liver (data not shown). Therefore, the apoC-I that is found in mammalian chylomicrons (34) must be acquired by these particles after they are secreted by the small intestine in the dog. In contrast, apoA-I mRNA is detected in both the liver and the small intestine, even though the concentration in the small intestine is only 15% of that in the liver (data not shown). As expected from its distribution in other mammals, apoE mRNA is present in a wide variety of tissues. By applying different amounts of RNA from these tissues and determining the linear regression coefficients (slopes) of each set of slot-blots, we calculated that the relative concentrations of apoE mRNA in

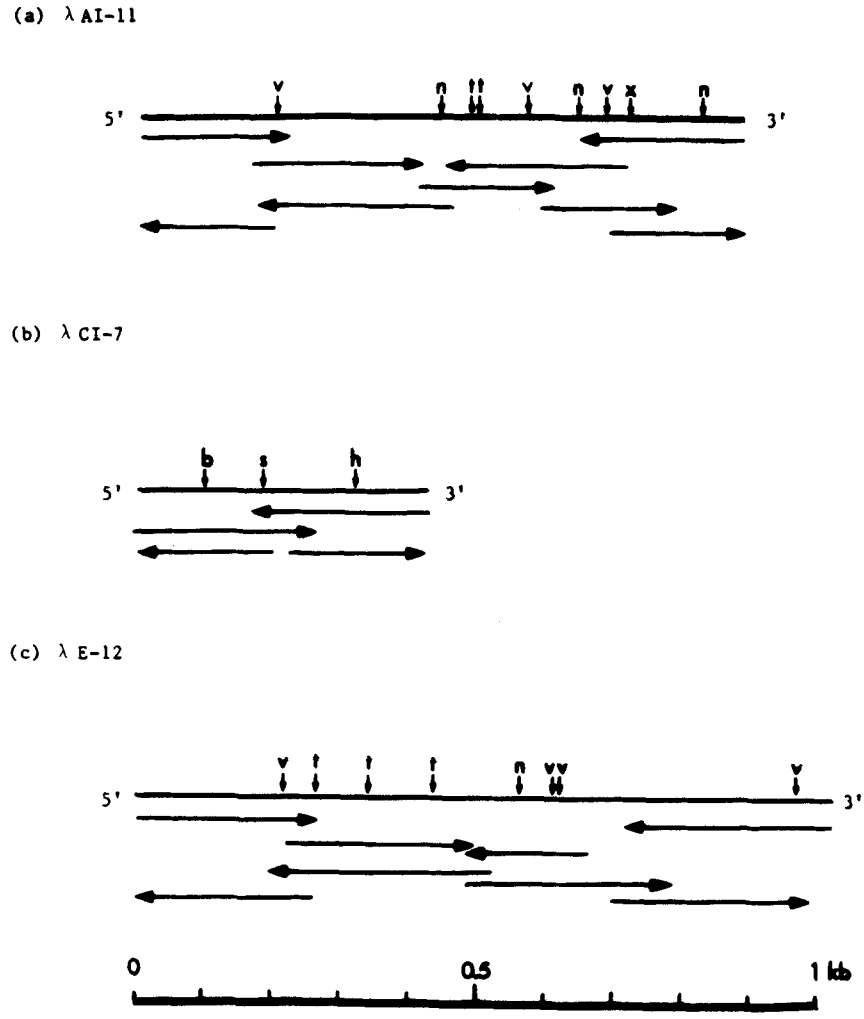

Fig. 1. Partial restriction maps and sequencing strategy of cloned dog apolipoprotein cDNAs. a) ApoA-I cDNA; b) apoC-I cDNA; and c) apoE cDNA. Sequencing was performed by the dideoxynucleotide chain termination technique of Sanger et al. (28). The direction and extent of each sequencing reaction are represented by the arrows. Synthetic oligonucleotides were used as sequencing primers. Restriction enzymes: b, BglI; h, HaeI; n, NarI; s, SmaI; t, PstI; v, PuvII; x, XhoI.

the tissues are 100, 17.5, 7.5, 6.9, 5.9, 5.5, 5.0, 3.3, 1.1, and 1.0, respectively, in the following organs: liver, jejunum, urinary bladder, ileum, colon, brain, kidney, spleen, pancreas, and testis (**Fig. 5**).

Rates of nucleotide substitution in the apoA-I, C-I, and E genes

First, we consider the rate of synonymous substitution, which can be computed from the number of substitutions per synonymous site (K_S) between genes (**Table 1**); apoC-I will not be considered because its coding region is short so that the estimate of K_S has a large standard error. The dog species is thought to have branched off slightly earlier than the divergence among the human, rat, and rabbit species (35) and may therefore be used as a reference to infer the K_S values in the latter lineages. Let a and b be the lengths from the ancestral node of the human and rat lineages to human and rat, and c be the length from the same node to dog (we consider the rabbit species below because its evolutionary position is uncertain). As calculated in **Table 2**, a = 0.20, b = 0.53, and c = 0.13 for apoA-I and a = 0.12, b = 0.47, and c = 0.17 for apoE, the averages being a = 0.16, b = 0.50, and c = 0.15. Therefore, the synonymous rate in the rat lineage is approximately three times higher than those in the human and dog lineages. The K_s values in the dog, human, and rabbit lineages (denoted by a, c, and d, respectively) are estimated to be a = 0.17, c = 0.14, and d = 0.15, suggesting that the synonymous rates in these three species are similar.

Next, we consider the rate of nonsynonymous substitu-

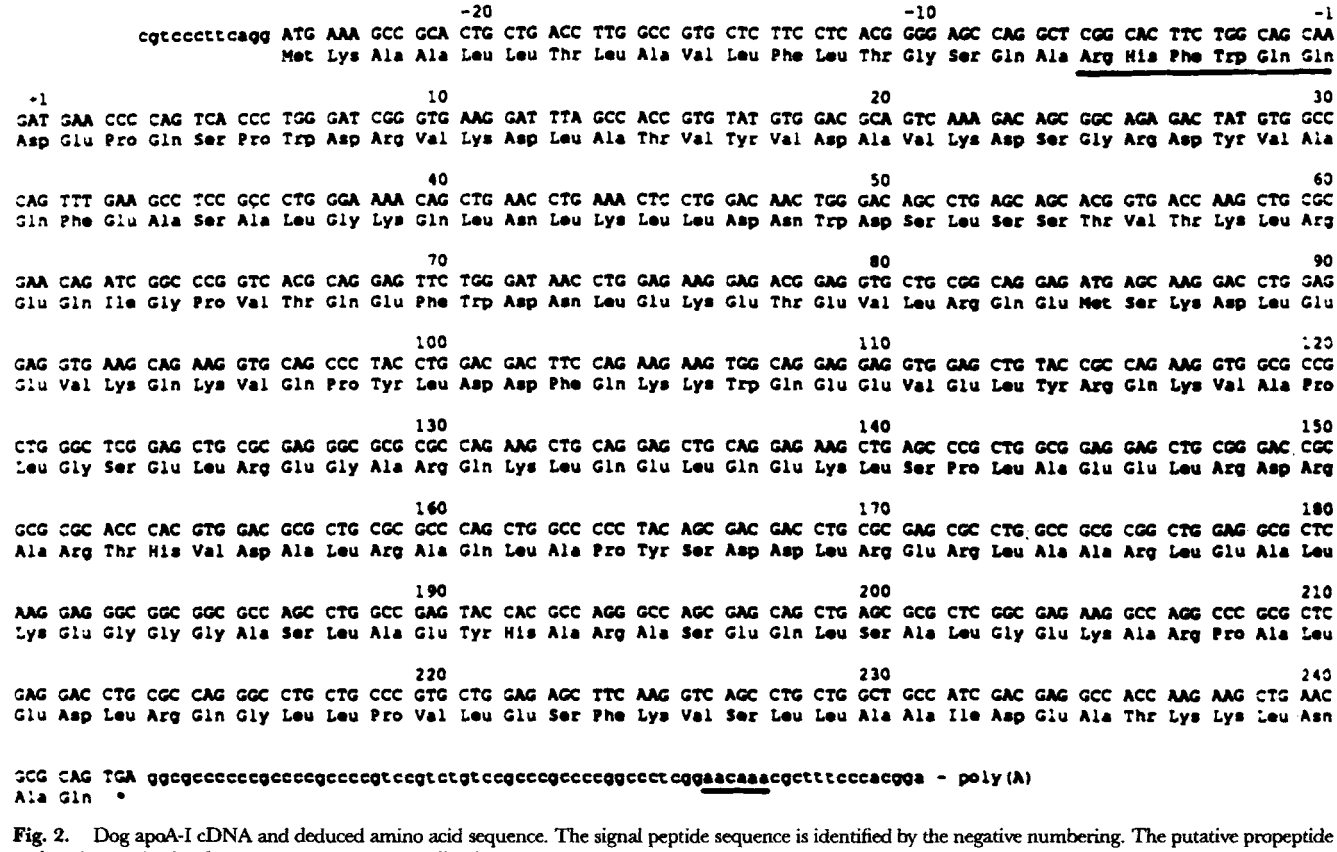

Fig. 2. Dog apoA-I cDNA and deduced amino acid sequence. The signal peptide sequence is identified by the negative numbering. The putative propeptide and variant polyadenylation sequences are underlined.

tion in the signal peptide region. in apoA-I, the K_A values for the human, rat, rabbit and dog lineages are 0.00, 0.16, 0.03, and 0.03, and in apoE, the K_A values for the human, rat, and rabbit lineages are 0.04, 0.18, and 0.22. These results suggest that very few nonsynonymous substitutions have occurred in the signal peptide region of the human apoA-I and apoE genes and in that of the dog apoA-I gene since the time of mammalian radiation. The same appears to be true for human and dog apoC-I because the K_A value in this region is only 0.06 between human and dog (Table 1). It is not clear why the signal peptides of these proteins in the human and dog lineages have been so well conserved. In the rabbit lineage, the signal peptide in apoA-I has evolved at a low rate (0.38×10^{-9}), whereas that in the apoE has evolved at a very high rate (2.75×10^{-9}). The signal peptide in rat apoA-I and apoE have evolved at a high

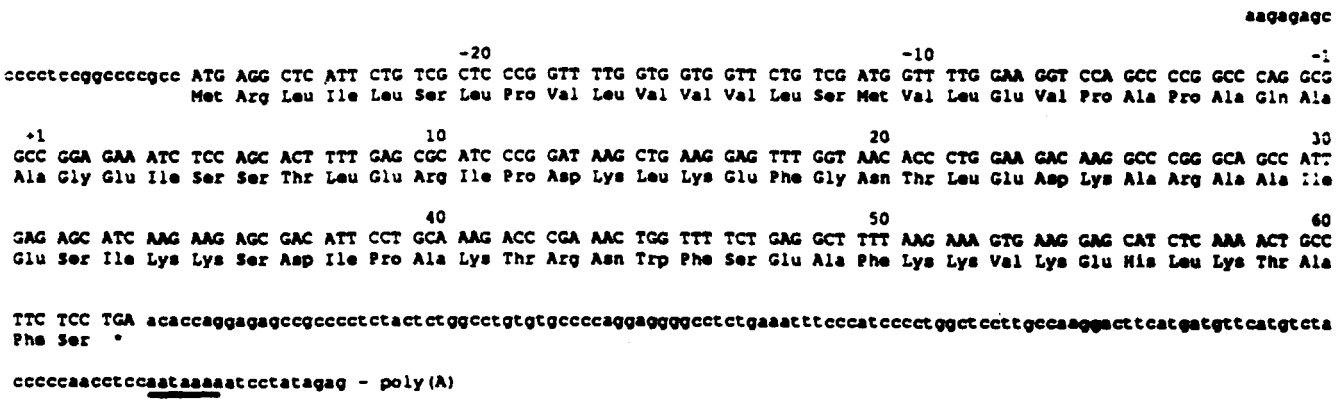

Fig. 3. Dog apoC-I cDNA and deduced amino acid sequence. The signal peptide sequence is identified by the negative numbering. The polyadenylation sequence is underlined.

```
 -1                                  10                                  20                                  30
AAG GTC CAG CAG GAG CTG GAG CCA GAG GCC GGG TGG CAG ACT GGC CAG CCC TGG GAG GCG GCG CTG GCC CGC TTC TGG GAT TAC CTG CGC
Lys Val Gln Gln Glu Leu Glu Pro Glu Ala Gly Trp Gln Thr Gly Gln Pro Trp Glu Ala Ala Leu Ala Arg Phe Trp Asp Tyr Leu Arg

                                    40                                  50                                  60
TGG GTG CAG ACG CTG TCT GAC CAG GTG CAA GAG GGC GTG CTC AAC ACC CAG GTC ACC CAG GAA CTG ACG GCG CTG ATG GAT GAG ACC ATG
Trp Val Gln Thr Leu Ser Asp Gln Val Gln Glu Gly Val Leu Asn Thr Gln Val Thr Gln Glu Leu Thr Ala Leu Met Asp Glu Thr Met

                                    70                                  80                                  90
AAG GAG GTG AAG GCC TAC AAG GCG GAG CTG GAC GAG CAG CTG GGC CCC ATG ACC TCG GAG ACG CAG GCC CGC GTG GCC AAG GAG CTG CAG
Lys Glu Val Lys Ala Tyr Lys Ala Glu Leu Asp Glu Gln Leu Gly Pro Met Thr Ser Glu Thr Gln Ala Arg Val Ala Lys Glu Leu Gln

                                   100                                 110                                 120
GCG GCG CAG GCC CGG CTG CGT GCG GAC ATG GAG GAC GTG CGC AAC CGC CTG ACG CAG TAC CGC GGC GAG CTG CAG GCC ATG CTG GGC CAG
Ala Ala Gln Ala Arg Leu Arg Ala Asp Met Glu Asp Val Arg Asn Arg Leu Thr Gln Tyr Arg Gly Glu Leu Gln Ala Met Leu Gly Gln

                                   130                                 140                                 150
AGC AGC GAG GAG CTC CGG GCG CGC TTC GCC TCC CAC ATG CGC AAG TTG CGT AAG CGG GTG CTG CGG GAC GCC GAG GAC CTG CAG AGG CGC
Ser Ser Glu Glu Leu Arg Ala Arg Phe Ala Ser His Met Arg Lys Leu Arg Lys Arg Val Leu Arg Asp Ala Glu Asp Leu Gln Arg Arg

                                   160                                 170                                 180
CTG GCC GTC TAC AAG GCC GGC GTG CGC GAG GGT GCC GAG CGC AGC GTG AGC AGC ATC CGC GAG CGC CTC TGG CCG CTG CTG GAG CAG GCC
Leu Ala Val Tyr Lys Ala Gly Val Arg Glu Gly Ala Glu Arg Ser Val Ser Ser Ile Arg Glu Arg Leu Trp Pro Leu Leu Glu Gln Ala

                                   190                                 200                                 210
CGC GAG CGC AAC GCC AAG GTG GGC GCC CTG GCC ACG CAG CCG CTG CTC GAG CGG GCC GAC GCC TGG GGC CAG CAG CTG CGC GGC CAG CTG
Arg Glu Arg Asn Ala Lys Val Gly Ala Leu Ala Thr Gln Pro Leu Leu Glu Arg Ala Asp Ala Trp Gly Gln Gln Leu Arg Gly Gln Leu

                                   220                                 230                                 240
GAG GAG ATG AGC AGC CGG GCC CGC GGC CAC CTG GAG GAG ATG CGC GAG CAG ATA CAG GAG GTG CGG GTG AAG ATG GAG GAG CAG GCC GAC
Glu Glu Met Ser Ser Arg Ala Arg Gly His Leu Glu Glu Met Arg Glu Gln Ile Gln Glu Val Arg Val Lys Met Glu Glu Gln Ala Asp

                                   250                                 260                                 270
CAG ATA CGC CAA AAG GCC GAG GCC TTC CAG GCG CGC CTC AAG AGC TGG TTC GAG CCC CTG CTG GAA GAC ATG CAG CGC CAG TGG GAC GGG
Gln Ile Arg Gln Lys Ala Glu Ala Phe Gln Ala Arg Leu Lys Ser Trp Phe Glu Pro Leu Leu Glu Asp Met Gln Arg Gln Trp Asp Gly

                                   280                                 290
CTG GTG GAG AAG GTG CAG GCG GCC GTG GCC ACC ATC CCC ACC TCT AAG CCT GTG GAG GAA CCA TGA gcgccccgcatgccacctgctgggcctcccc
Leu Val Glu Lys Val Gln Ala Ala Val Ala Thr Ile Pro Thr Ser Lys Pro Val Glu Glu Pro  *

ccactccttccccgtgccccctgcctgctcccacgcctccaggaggctgccccctgccccccagctgtcctcctgaaaggggccctagcttaataaagattcaccaaagctccaccc - poly(A)
```

Fig. 4. Dog apoE cDNA and deduced amino acid sequence. The polyadenylation sequence is underlined.

rate, approximately 2.0×10^{-9} substitutions per site per year, which is two times the average rate (0.9×10^{-9}) for human and rat genes (29).

Finally, we consider the nonsynonymous rate in the mature peptide region. For the divergence between dog and human apoC-I, the nonsynonymous rate is 1.4×10^{-9}, which is considerably higher than the average rate for human and rat genes (29). In apoA-I, the K_A values are approximately 0.055, 0.195, 0.065, and 0.042 for the human, rat, rabbit, and dog lineages and the corresponding nonsynonymous rates are 0.66×10^{-9}, 2.44×10^{-9}, 0.81×10^{-9}, and 0.53×10^{-9}. In apoE, the K_A values are approximately 0.055, 0.13, 0.09, and 0.11 for the human, rat, rabbit, and dog lineages, the corresponding rates being 0.69×10^{-9}, 1.63×10^{-9}, 1.13×10^{-9}, and 1.38×10^{-9}. In both apoA-I and apoE, the rate in the rat lineage is three to four times the rate in the human lineage. Previously, from a comparison of human and rat apoA-I and apoE genes, we concluded that apoA-I has evolved considerably faster than apoE (36). However, it is now clear from the above computation that this is true only for the rat lineage, though it is not clear why rat apoA-I has evolved exceptionally fast. In the human lineage, apoA-I and apoE have evolved at the same rate,

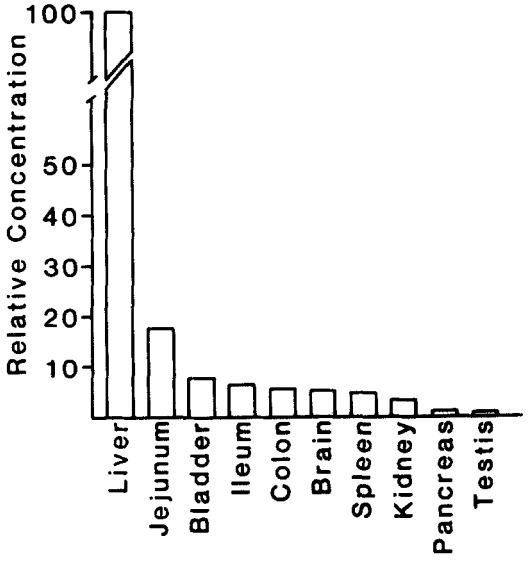

Fig. 5. Relative concentration of apoE mRNA in dog tissues. The concentrations are deduced from the regression coefficients (slopes) of densitometric measurements of autoradiographs of slot blots obtained from graded amounts of dog total RNA hybridized to ^{32}P-labeled dog apoE cDNA clone, λE-12.

TABLE 1. Number of nucleotide substitutions per synonymous site (K_S) and per nonsynonymous site (K_A) between mammalian apolipoprotein genes

Gene	Species Pair	K_S	K_A	
			Signal Peptide	Mature Peptide
C-I	Human vs dog	0.53 ± 0.13	0.06 ± 0.03	0.23 ± 0.05
A-I	Human vs dog	0.33 ± 0.05	0.03 ± 0.03	0.09 ± 0.01
A-I	Human vs rat	0.73 ± 0.10	0.15 ± 0.07	0.26 ± 0.02
A-I	Human vs rabbit	0.34 ± 0.05	0.03 ± 0.03	0.12 ± 0.02
A-I	Dog vs rat	0.66 ± 0.09	0.19 ± 0.08	0.24 ± 0.02
A-I	Dog vs rabbit	0.32 ± 0.05	0.06 ± 0.04	0.10 ± 0.01
A-I	Rat vs rabbit	0.68 ± 0.09	0.19 ± 0.08	0.25 ± 0.02
E	Human vs dog	0.28 ± 0.04	NA	0.16 ± 0.02
E	Human vs rat	0.59 ± 0.07	0.22 ± 0.09	0.18 ± 0.02
E	Human vs rabbit	0.34 ± 0.05	0.26 ± 0.10	0.14 ± 0.02
E	Dog vs rat	0.64 ± 0.08	NA	0.23 ± 0.02
E	Dog vs rabbit	0.29 ± 0.04	NA	0.19 ± 0.02
E	Rat vs. rabbit	0.69 ± 0.09	0.40 ± 0.13	0.22 ± 0.02

whereas in the rabbit and dog lineages, apoE has evolved faster than apoA-I. Therefore, on the average, apoA-I seems to be more conservative than apoE.

From a comparison of a partial rat apoB cDNA sequence (37) corresponding to residues 595 to 979 in human apoB with the human apoB cDNA sequence (38), we obtain $K_A = 0.09 \pm 0.01$ and a nonsynonymous rate of 0.56×10^{-9}. As this rate is considerably lower than the average rates for human and rat apoA-I and apoE, this part of apoB is probably more conservative than apoA-I and apoE. A partial sequence for the 3' end of the chicken apoB cDNA is now available (39) and we have estimated that the K_A value between this sequence and the corresponding human sequence (codons for residues 4204–4536) is 0.59 ± 0.04. Assuming 300 million years for the divergence between mammals and birds, we obtain a rate of 0.98×10^{-9}. On the other hand, the K_A value between chicken and human apoA-I (40,41) is 0.41 ± 0.03, corresponding to a rate of 0.68×10^{-9}. Thus, this part of the apoB appears to be less conservative than apoA-I.

Comparison of protein sequences

We have aligned, for each protein, the sequences from different species (**Fig. 6**). In each of these three proteins, as well as the other soluble apolipoproteins, there is a common block of 33 residues at the end of exon 3 (42), and the region encoded by exon 4 contains repeats of 11 or 22 residues, which are labeled as repeats A-I-4, A-I-5, etc.

Although apoA-I has evolved rather rapidly, many residues have been conserved among all the five vertebrate species (indicated by #), or among all the four mammalian species (indicated by +) (Fig. 6). Further, a close examination reveals that most of the amino acid substitutions occur between residues of similar biochemical properties, e.g., hydrophobicity. We noted above that rat apoA-I has evolved exceptionally fast in terms of nucleotide substitutions. Fig. 6A shows that this is also true in terms of deletions: rat apoA-I contains four deletions in the boxed region, whereas among the other four sequences, only rabbit apoA-I contains a deletion in this region.

TABLE 2. Rate of substitution per site per year

Gene	Branch Length	From K_S Values	Rate	From K_A Values			
				Signal Peptide	Rate	Mature Peptide	Rate
A-I	a	0.20	2.5×10^{-9}	0.00	0.00×10^{-9}	0.055	0.69×10^{-9}
A-I	b	0.53	6.6×10^{-9}	0.16	2.00×10^{-9}	0.195	2.44×10^{-9}
A-I	c	0.13	1.6×10^{-9}	0.03	0.38×10^{-9}	0.042	0.53×10^{-9}
E	a	0.16	2.0×10^{-9}	NA	NA	0.055	0.69×10^{-9}
E	b	0.50	6.3×10^{-9}	NA	NA	0.130	1.63×10^{-9}
E	c	0.15	1.9×10^{-9}	NA	NA	0.110	1.38×10^{-9}

The parameters a, b, and c are the lengths from the node connecting the human, rat, and dog lineages to human (H), rat (R), and dog (D). They are given by $a = (D_{HR} + D_{HD} - D_{RD}/2)$, $b = (D_{HR} + D_{RD} - D_{HD}/2)$ and $c = (D_{HD} + D_{RD} - D_{HR}/2)$, where D_{XY} is the distance (K_S or K_A) between species X and Y; for the method, see Fitch and Margoliash (60). The substitution rates are calculated under the assumption that the three species diverged 80 million years ago (61). NA: The nucleotide sequence for the signal peptide of dog apoE is not available.

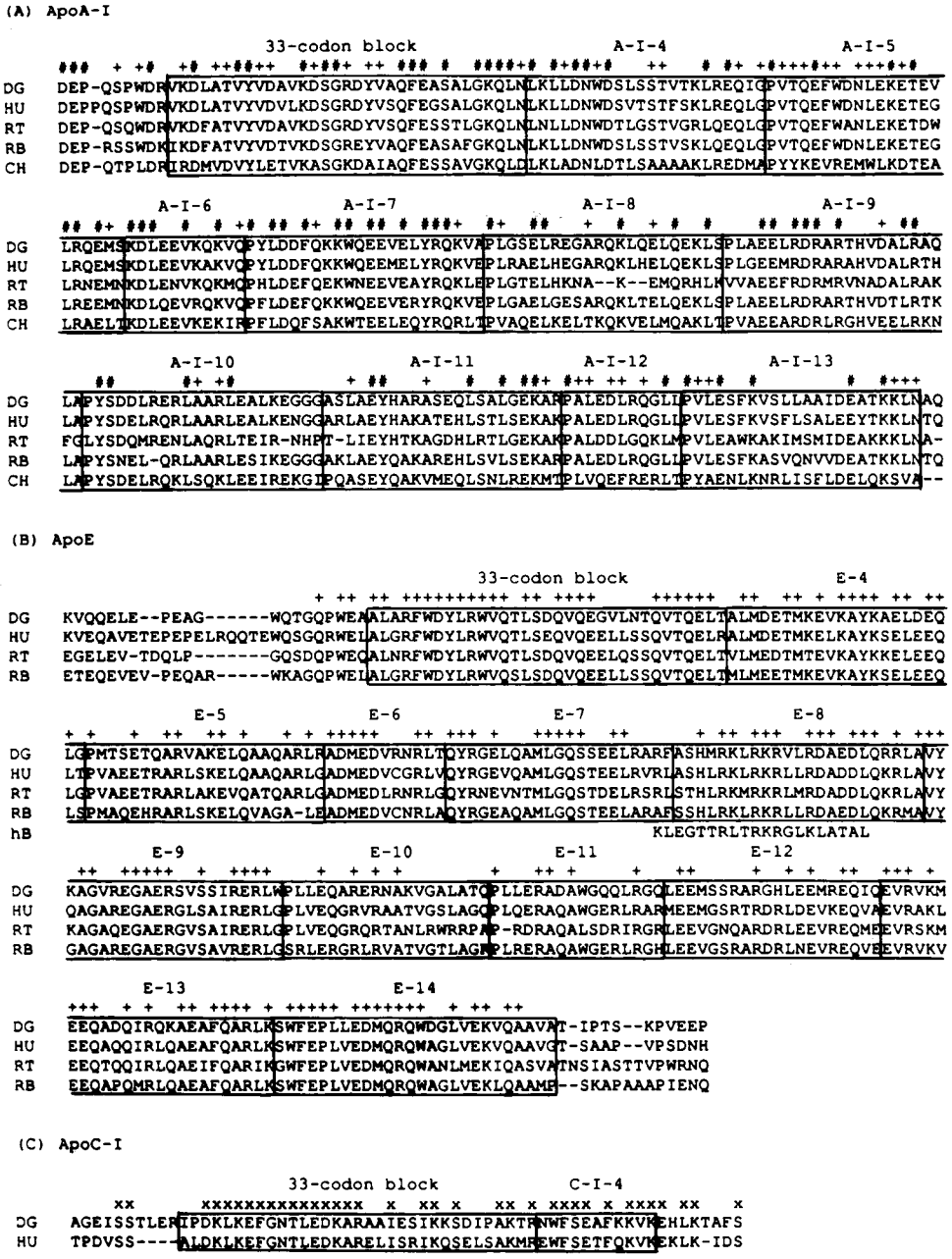

Fig. 6. Comparison of amino acid sequences. In each protein, there is a common block of 33 residues at the end of exon 3 (42), and the region encoded by exon 4 contains repeats of 11 or 22 residues, which are labeled as repeats A-I-4, A-I-5, etc.; all these regions are boxed. (A) ApoA-I. The residues that are conserved among all the five vertebrate species are indicated by # and those that are conserved among the four mammalian species but not in chicken are indicated by +. (B) ApoE. The residues that are conserved among the four mammalian species are indicated by +. (C) ApoC-I. The residues conserved in the two mammalian species compared are indicated by x. Species notation: DG, dog; HU, human; RT, rat; RB, rabbit; CH, chicken; hB, human apoB putative receptor binding region (54, 55).

The above conclusion that apoE is less conservative than apoA-I is supported by the observation from Fig. 6A and B that fewer residues in apoE have been conserved among the four mammalian species and by the fact that the repeat pattern is better conserved in apoA-I than in apoE (42).

For apoC-I, only two sequences are now available so it is difficult to identify nonconserved regions. However, it is clear from Fig. 6 that all regions of this protein are less conserved than apoA-I and apoE.

We note that, in each protein, the degree of sequence conservation varies among regions. For example, the common block of 33 residues has been well conserved, whereas the unboxed regions contain many changes including deletions and insertions. The implications of conservation of in-

dividual regions in each protein for structure–function will be discussed later.

Origin of the human apoC-I pseudogene

We have compared the human apoC-I pseudogene (43) with its functional counterpart. In the coding region, the K_S and K_A values are, respectively, 0.07 ± 0.04 and 0.13 ± 0.03, the latter being almost twice the former. This is surprising because, in a functional gene, K_S is usually much larger than K_A (see Tables 1 and 2). Although the low K_S value is probably an extreme random deviate, because the region compared is short, this observation suggests that the pseudogene became nonfunctional soon after its formation; after that nonsynonymous substitutions were not subject to functional constraints and could occur at a high rate. This suggestion is supported by the fact that the two genes have diverged almost as fast as Alu sequences, which are now commonly thought to be pseudogenes (44); for the five Alu sequences at corresponding positions in the two genes (43), the numbers of nucleotide substitutions per site are 0.13, 0.15, 0.13, 0.16, and 0.16, which are only slightly higher than the K_A value given above. The average for the five numbers is 0.15, which is higher than the number (0.11) of substitutions per site between the human and owl monkey η globin pseudogene (45). Thus, the apoC-I pseudogene probably arose earlier than the divergence between the human and owl (New World) monkey lineages, i.e., about 35 million years ago. It will be interesting to see whether this pseudogene is in fact also present in New World monkeys.

DISCUSSION

Tissue expression of dog apolipoprotein mRNAs

ApoC-I is expressed only in the liver, whereas apoA-I is expressed both in liver and in the intestine, even though the apoA-I mRNA concentration in the latter tissue is only ~ 15% of that in the liver. In contrast, apoE mRNA is present in many tissues including liver, small intestine, urinary bladder, colon, brain, kidney, spleen, pancreas, and testis. The widespread distribution of apoE synthesis is consistent with the experiments in vivo showing that newly synthesized apoE present in interstitial fluids contributes substantially to the plasma apoE pool in dogs (46). The almost ubiquitous presence of apoE mRNA in the dog is reminiscent of similar findings in other mammals, including humans, nonhuman primates, and rodents (16–20). Like other mammals, canine brain tissue contains a substantial amount of apoE mRNA. It is the source of the fairly high concentration of cerebrospinal fluid apoE in the dog (38–60% of the plasma concentration compared to ~ 5.4% in man) (47). In other mammals, the small intestine expresses little, if any, apoE mRNA. In contrast, dog small intestine seems to contain a substantial amount of the mRNA, suggesting that in this species, the small intestine is potentially a source of an appreciable amount of circulating apoE. Since the RNA was prepared from total jejunum, we cannot be certain whether the mRNA is derived from the mucosa or submucosal tissues or both.

To date, we have examined the tissue distribution of a number of dog apolipoprotein mRNAs, including those of apoC-II, C-III (36), C-I, A-I, and E (present study). These apolipoproteins showed substantial differences in their site of synthesis: one of them is expressed exclusively in the liver (apoC-I), one in both liver and intestine (apoA-I), and three others (apoC-II, C-III, and E) are expressed in a wide variety of tissues. Future studies using the dog as an experimental model for lipoprotein metabolism should take into consideration the relative tissue distribution of these mRNAs. The apolipoproteins synthesized in such tissues may perform some specific function locally. They also contribute to the circulating pool of apolipoproteins in this animal.

Relative rates of evolution and functional aspects of apolipoprotein structure

In apoA-I, the 33 residue common block is well conserved among species and so are repeats 4, 5, 6, and 7 (in each region, over 55% of the residues are conserved among the four mammalian species) (Fig. 6A), suggesting that the stringency of structural requirements in these regions is fairly strong. A major function of apoA-I is the activation of LCAT (10–12). Soutar et al. (23) have shown that both the amino- and carboxyl-terminal cyanogen bromide fragments (residues 1–85 and 147–243) activate LCAT; in the latter fragment, residues 145–182 seem to be involved in the activation process. 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 (Fig. 6A). Thus, the structural requirements for LCAT activation may not be stringent. This conclusion is supported by the fact that apoC-I, which can also activate LCAT, has evolved rapidly (see above). Moreover, synthetic model peptides that mimic apoA-I surface properties but differ from apoA-I in primary sequences are effective in LCAT activation (48). While LCAT activation is 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 its function.

All regions of apoC-I are less conserved than apoA-I and apoE. Thus, the functions of apoC-I, which have been suggested to be phospholipid binding and LCAT activation, do not have very stringent structural requirements.

In apoE, comparison of the internally repeated regions (the common block, and repeats E-4 to E-14, boxed in Fig. 6B), and the flanking nonrepeated sequences in the mature peptide (unboxed in Fig. 6B) indicates that the former are much better conserved than the latter. It suggests that the

repeats serve functions that have a more stringent structural requirement. In particular, the 33-codon common block is especially well conserved, whereas the region immediately preceding it has evolved much faster. Repeats 4, 6, 8, 9, 13, and 14 are better conserved than the other repeats.

In aqueous solution, there is evidence that apoE contains two independently folded structural domains: a relatively unstable self-associating carboxyl-terminal domain (residues 225–299) rich in amphipathic helices, and a more stable amino-terminal domain (residues 20–165) that resembles a soluble globular protein in structure (49, 50). The two domains are connected by an exposed peptide segment or hinge region that appears to have random coil structure and is highly susceptible to proteolysis. Inspection of Fig. 6B indicates that the amino- and carboxyl-terminal domains are considerably better conserved than the connecting hinge region. This variation in interspecies homology over the entire length of apoE supports the thesis that the two structural domains require much more stringent sequence conservation for their functions than does the hinge region.

ApoE is an important determinant in the interaction between apoE-containing lipoproteins and cell-surface receptors (51). Studies using monoclonal antibodies, natural mutants, and site-specific mutants produced in vitro localized the receptor-binding region to the vicinity of residues 140–150 and have thus far identified at least eight specific residues (#136, 140, 142, 143, 145, 146, 150, and 158) as crucial residues involved in receptor binding (52, 53). Further, the α-helical conformation in this region also appears essential, since substitution of Pro for either Leu^{144} or Ala^{152} will interfere with binding activity. Examination of the degree of conservation of the various repeats in apoE (Fig. 6B) indicates that E-8 (residues 130–166), which encompasses the receptor-binding region, is indeed one of the most highly conserved regions of apoE.

When we specifically compared the eight residues that have been directly implicated in receptor binding, only one substitution (an Ala for Arg^{136}, human equivalent) in the rabbit sequence was seen among the four mammalian sequences. In this case, the Ala was immediately preceded by an Arg not present in the other sequences. It is noteworthy that a subject heterozygous for apoE-2-Christchurch (resulting from an $Arg^{136}\rightarrow$Ser mutation) and classical apoE2 (i.e., $Arg^{145}\rightarrow$Cys) presented with Type III hyperlipoproteinemia (53). A genetically engineered $Arg^{136}\rightarrow$Ser mutant also showed only 41% of the normal apoE receptor binding to the LDL receptor (54). The relative receptor binding activity of rabbit apoE, which has an $Arg^{136}\rightarrow$Ala substitution, has not been determined. When we aligned the putative receptor binding sequence for human apoB-100 (residues 3352–3372) (designated hB in Fig. 6B, regs. 54, 55) with the corresponding apoE sequences, additional differences involving residues 136 (Arg→Lys), 140 (His→Thr), 143 (Lys→Leu), 150 (Arg→Lys), and 158 (Arg→Ser) are evident. This analysis suggests that though apoB-100 and apoE appear to bind to the same receptor, they might do so by interactions thay may not be identical, perhaps accounting for the significant differences in affinity between them (56). Furthermore, structure-function correlation studies on the human low density receptor support this interpretation. A mutant low density lipoprotein receptor from a patient with familial hypercholesterolemia has been described that has lost its binding affinity for low density lipoproteins, but has retained its ability to bind apoE (57). Finally, when different regions in the ligand binding domain of a cloned low density lipoprotein receptor are deleted, binding of the receptor to low density lipoproteins or to β-migrating very low density lipoproteins (containing apoE) is affected in a dissimilar manner (58). Therefore, apoB-100 and apoE interact with the low density lipoprotein receptor via distinct mechanisms. We had difficulty aligning the apoE sequences with the other putative receptor-binding region in apoB-100 (residues 3147–3157) proposed by Knott et al. (59). To date, experimental support for receptor-binding activity is not available for this domain. Our comparison suggests that if this domain is a bona fide receptor-binding sequence, its mechanism of interaction must differ even more in detail from the apoE-receptor interaction than does the domain located closer to the carboxyl terminus (residues 3352–3372).

This work was supported by National Institutes of Health Grants GM-39927 (to W. H. L.) and HL-16512 (to L. C.).

Manuscript received 20 March 1989 and in revised form 30 May 1989.

REFERENCES

1. Mahley, R.W., T. L. Innerarity, S. C. Rall, Jr., and K. H. Weisgraber. 1984. Plasma lipoprotein: apolipoprotein structure and function. *J. Lipid Res.* **25:** 1277–1294.
2. Li, W-H., M. Tanimura, C-C. Luo, S. Datta, and L. Chan. 1988. The apolipoprotein multigene family: biosynthesis, structure, structure-function relationships, and evolution. *J. Lipid Res.* **29:** 245–271.
3. Mahley, R. W. 1978. Alterations in plasma lipoproteins induced by cholesterol feeding in animals including man. *In* Disturbances in Lipid and Lipoprotein Metabolism. J. M. Dietschy, A. M. Gotto, Jr., and J. A. Ontko, editors. American Physiological Society, Bethesda, MD. 181–197.
4. Chuang, M. Y., L. Wong, W. R. Gallaher, J. J. Thompson, and P. S. Roheim. 1985. Production and characterization of a monoclonal antibody to dog hepatic lipase. *Biochim. Biophys. Acta.* **833:** 69–81.
5. Dory, L., L. M. Boquet, R. L. Hamilton, C. H. Sloop, and P. S. Roheim. 1985. Heterogeneity of dog interstitial fluid (peripheral lymph) high density lipoproteins: implications for a role in reverse cholesterol transport. *J. Lipid Res.* **26:** 519–527.
6. Barr, D. P., E. M. Russ, and H. A. Eder. 1951. Protein-lipid relationship in human plasma. II. In arteriosclerosis and related conditions. *Am. J. Med.* **11:** 480–493.
7. Gofman, J. W., O. DeLalla, F. Glazier, N. K. Freeman, F. T. Lindgren, A. V. Nichols, E. H. Strisower, and A. R. Tamplin. 1954. The serum lipoprotein transport system in health, meta-

bolic disorders, atherosclerosis and coronary heart disease. *Plasma.* **2:** 413–484.
8. Miller, G. J., and N. E. Miller. 1975. Plasma high density lipoprotein concentration and devleopment of ischaemic heart disease. *Lancet.* **1:** 16–19.
9. Gordon, T., W. P. Castelli, M. C. Hortland, W. B. Kannel, and T. R. Dawber. 1977. High density lipoproteins as a protective factor against coronary heart disease. *Am. J. Med.* **62:** 707–714.
10. Glomset, J. A. 1968. The plasma lecithin:cholesterol acyltransferase reaction. *J. Lipid Res.* **9:** 155–167.
11. 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.
12. Fielding, C. J., V. G. Shore and P. E. Fielding. 1972. Lecithin:cholesterol acyltransferase: effects of substrate composition upon enzyme activity. *Biochim. Biophys. Acta.* **270:** 513–518.
13. Brown, M. S., P. T. Kovanen, and J. L. Goldstein. 1981. Regulation of plasma cholesterol by lipoprotein receptors. *Science.* **212:** 628–635.
14. 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.
15. Mahley, R. W., T. L. Innerarity, and K. H. Weisgraber. 1980. Alterations in metabolic activity of plasma lipoproteins following selective chemical modification of the apoproteins. *Ann. N. Y. Acad. Sci.* **348:** 265–280.
16. Blue, M-L., D. L. Williams, S. Zucker, S. A. Khan, and C. B. Blum. 1983. Apolipoprotein E synthesis in human kidney, adrenal gland, and liver. *Proc. Natl. Acad. Sci. USA.* **80:** 283–287.
17. Elshourbagy, N. S., W. S. Liao, R. W. Mahley, and J. M. Taylor. 1984. Apolipoprotein E mRNA is abundant in the brain and adrenals, as well as in the liver, and is present in other peripheral tissues of rats and marmosets. *Proc. Natl. Acad. Sci. USA. USA.* **82:** 203–207.
18. Lin-Lee, Y.C., F. T. Kao, P. Cheung, and L. Chan. 1985. Apolipoprotein E gene mapping and expression: localization of the mRNA in lipoprotein and non-lipoprotein producing tissues. *Biochemistry.* **24:** 3751–3756.
19. Lin, C-T., Y. Xu, J-Y. Wu, and L. Chan. 1986. Immunoreactive apolipoprotein E is a widely distributed cellular protein. Immunohistochemical localization of apolipoprotein E in baboon tissues. *J. Clin. Invest.* **78:** 947–958.
20. Ignatius, M. J., E. M. Shooter, R. E. Pitas, and R. W. Mahley. 1987. Lipoprotein uptake by neuronal growth cones in vitro. *Science.* **236:** 959–962.
21. Avila, E. M., G. Holdsworth, M. Sasaki, R. L. Jackson, and J. A. Harmony. 1982. Apoprotein E suppresses phytohemagglutinin-activated phospholipid turnover in peripheral blood mononuclear cells. *J. Biol. Chem.* **257:** 5900–5909.
22. Mahley, R. W. 1988. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science.* **240:** 522–530.
23. 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.
24. Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* **18:** 5294–5299.
25. Aviv, H., and P. Leder. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. USA.* **69:** 1408–1412.
26. Gubler, U., and B. J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene.* **25:** 263–269.
27. Maniatis, T., Z. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
28. Sanger, F., A. Coulson, B. Barrell, A. Smith, and B. Roe. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* **143:** 161–178.
29. Li, W-H., C-I. Wu, and C-C. Luo. 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.
30. Gordon, J. I., H. F. Sims, S. R. Lentz, C. Edelstein, A. M. Scanu, and A. W. Strauss. 1983. Proteolytic processing of human preproapolipoprotein A-I: a proposed defect in the conversion of proA-I to A-I in Tangier disease. *J. Biol. Chem.* **258:** 4037–4044.
31. 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 proprotein. *Proc. Natl. Acad. Sci. USA.* **80:** 2574–2478.
32. Chung, H., A. Randolph, I. Reardon, and R. L. Heinrickson. 1982. The covalent structure of apolipoprotein A-I from canine high density lipoproteins. *J. Biol. Chem.* **257:** 2961–2967.
33. Weisgraber, K. H., R. F. Troxler, S. C. Rall, and R. W. Mahley. 1980. Comparison of the human, canine and swine E apoproteins. *Biochem. Biophys. Res. Commun.* **95:** 374–380.
34. Wu, A-L., and H. G. Windmueller. 1979. Relative contributions by liver and intestine to individual plasma apolipoproteins. *J. Biol. Chem.* **254:** 7316–7322.
35. Dayhoff, M. O. 1972. Atlas of Protein Sequences and Structure. Vol. 5. National Biomedical Research Foundation, Silver Spring, MD.
36. 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.
37. Matsumoto, A., H. Aburatani, Y. Shibasaki, T. Kodamn, F. Takake, and H. Itakura. 1987. Cloning and regulation of rat apolipoprotein B mRNA. *Biochem. Biophys. Res. Commun.* **142:** 92–99.
38. 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.
39. Kirchgessner, T. G., C. Heinzmann, K. L. Svenson, D. A. Gordon, M. Nicosia, H. G. Lebherz, A. J. Lusis, and D. L. Williams. 1987. Regulation of chicken apolipoprotein B: cloning, tissue distribution, and estrogen induction of mRNA. *Gene.* **59:** 241–251.
40. Cheung, P., and L. Chan. 1983. Nucleotide sequence of cloned cDNA of human apolipoprotein A-I. *Nucleic Acids Res.* **11:** 3703–3715.
41. 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.
42. 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.
43. Lauer, S. J., D. Walker, N. A. Elshourbagy, C. A. Reardon, B. Levy-Wilson, and J. M. Taylor. 1988. Two copies of the human apolipoprotein C-I gene are linked closely to the apolipoprotein E gene. *J. Biol. Chem.* **263:** 7277–7286.
44. Britten, R. J., W. F. Baron, D. B. Stout, and E. H. Davidson. 1988. Sources and evolution of human Alu repeated sequences. *Proc. Natl. Acad. Sci. USA.* **85:** 4770–4774.
45. Li, W-H., M. Tanimura, and P. M. Sharp. 1987. An evaluation of the molecular clock hypothesis using mammalian DNA sequences. *J. Mol. Evol.* **25:** 330–342.
46. Dory, L., L. M. Boquet, C. R. Tate, and C. H. Sloop. 1986. Peripheral synthesis and isoform distribution of dog apoprotein E: an in vivo approach. *J. Biol. Chem.* **261:** 811–816.
47. Pitas, R. E., J. K. Boyles, S. H. Lee, D. Hui, and K. H. Weisgraber. 1987. Lipoproteins and their receptors in the central nervous system. Characterization of the lipoproteins in cerebrospinal fluid and identification of apolipoprotein B, E (LDL) receptors in the brain. *J. Biol. Chem.* **262:** 14352–14360.
48. 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.
49. Wetterau, J. R., L. P. Aggerbeck, S. C. Rall, Jr., and K. H. Weisgraber. 1988. Human apolipoprotein E3 in aqueous solution. I. Guidance for two structural domains. *J. Biol. Chem.* **263:** 6240–6248.
50. Aggerbeck, L. P., J. R. Wetterau, K. H. Weisgraber, C-S. C. Wu, and F. T. Lindgren. 1988. Human apolipoprotein E3 in aqueous solution. II. Properties of the amino- and carboxyl-terminal domains. *J. Biol. Chem.* **263:** 6249–6258.
51. Mahley, R. W., and T. L. Innerarity. 1983. Lipoprotein receptors and cholesterol homeostasis. *Biochim. Biophys. Acta.* **737:** 197–222.
52. Lalazar, A., K. H. Weisgraber, S. C. Rall, Jr., H. Giladi, T. L. Innerarity, A. Z. Levanon, J. K. Boyles, B. Amit, M. Gorecki, R. W. Mahley, and T. Vogel. 1988. Site specific mutagenesis of human apolipoprotein E. Receptor binding activity of variants with single amino acid substitutions. *J. Biol. Chem.* **263:** 3542–3545.
53. Wardell, M. R., S. O. Brennan, E. D. James, R. Fraser, and R. W. Carrell. 1987. Apolipoprotein E2-Christchurch (136 Arg → Ser). New variant of human apolipoprotein E in a patient with Type III hyperlipoproteinemia. *J. Clin. Invest.* **80:** 483–490.
54. 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. Beyers, L. M. Priestley, E. Robertson, L. B. Rall, C. Betsholz, 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.
55. 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.
56. 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.
57. Hobbs, H. H., M. S. Brown, J. L. Goldstein, and D. W. Russell. 1986. Deletion of exon encoding cysteine-rich repeat of low density lipoprotein receptor alters its binding specificity in a subject with familial hypercholesterolemia. *J. Biol. Chem.* **261:** 13114– 13120.
58. Esser, V., L. E. Limbird, M. S. Brown, J. L. Goldstein and D. W. Russell. 1988. Mutational analysis of the ligand binding domain of low density lipoprotein receptor. *J. Biol Chem.* **263:** 13282–13290.
59. Knott, T. J., R. J. Pease, L. M. Powell, S. C. Wallis, S. C. Rall, T. L. Innerarity, B. Blackhart, W. H. Taylor, Y. 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.
60. 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:** 279–294.
61. Romer, A. S. 1966. Vertebrate Paleontology, University of Chicago Press, Chicago.