Registry No. DPPC, 2644-64-6; proton, 12586-59-3; dansylglycine, 1091-85-6.

REFERENCES

Bangham, A. D., de Gier, J., & Greville, G. D. (1967) *Chem. Phys. Lipids 1*, 225–237.

Bramhall, J. (1984) *Biochim. Biophys. Acta 778*, 393–399.

Bramhall, J. (1985) *AOCS Monogr. 13*, 43–50.

Bramhall, J. (1986) *Biochemistry* (in press).

Cafiso, D. S., & Hubbell, W. L. (1983) *Biophys. J. 44*, 49–57.

Deamer, D. W., Prince, R. C., & Crofts, A. R. (1972) *Biochim. Biophys. Acta 274*, 323–335.

Garcia, M. L., Kitada, M., Eisenstein, H. C., & Krulwich, T. A. (1984) *Biocheim. Biophys. Acta 766*, 109–115.

Nichols, J. W., Hill, M. W., Bangham, A. D., & Deamer, D. W. (1980) *Biochim. Biophys. Acta 596*, 393–403.

Nozaki, Y., & Tanford, C. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4324–4328.

Rossignol, M., Thomas, P., & Grignon, C. (1982) *Biochim. Biophys. Acta 684*, 195–199.

Rottenberg, H. (1975) *Bioenergetics 7*, 61–74.

Schullery, S. E., Schmidt, C. F., Felgner, P., Tillack, T. W., & Thompson, T. E. (1980) *Biochemistry 19*, 3919–3923.

Thomas, J. A., Buchsbaum, R. N., Zimniak, A., & Racker, E. (1979) *Biochemistry 18*, 2210–2218.

Waddell, W. J., & Butler, T. C. (1959) *J. Clin. Invest. 38*, 720–729.

Structure, Evolution, and Tissue-Specific Synthesis of Human Apolipoprotein AIV†

Sotirios K. Karathanasis* and Ivan Yunis

Laboratory of Molecular and Cellular Cardiology, Department of Cardiology, Children's Hospital/Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115

Vassilis I. Zannis

Section of Molecular Genetics, Cardiovascular Institute, Boston University Medical Center, Boston, Massachusetts 02118

Received November 13, 1985; Revised Manuscript Received February 25, 1986

ABSTRACT: Apolipoprotein AIV (apoAIV) is a protein of the lipid transport system found associated with chylomicrons, high-density lipoprotein (HDL), and the lipoprotein-free fraction of the plasma. The gene coding for the human apoAIV is closely linked with the genes coding for apolipoproteins AI (apoAI) and CIII (apoCIII). In this paper a nearly full-length apoAIV cDNA clone has been isolated by screening an adult human liver DNA library using a human apoAIV gene probe. In-frame translation of the cDNA sequence in this clone indicated that the human apoAIV consists of 396 amino acid residues including a 20 residue long signal peptide. The coding region of this cDNA sequence contains 15 nucleotide repeats, 11 of which code for amino acid repeats with potentials of forming amphipathic helices. Alignment and comparison of the human and rat apoAIV amino acid sequences indicated a five-residue deletion near the carboxy terminus of the rat protein. This comparison also indicated that these proteins are 61.8% homologous, suggesting that the rate of evolution of apoAIV is 65 accepted point mutations (PAMs) per 100 residues per 100 million years. The rates of evolution of certain amino acid repeats in apoAIV are higher than the rate of evolution of the entire protein. However, the corresponding, computer-generated, secondary structures and hydropathy profiles of these repeats are very similar between the human and rat apoAIV. The relative steady-state levels of apoAIV mRNA in various human and monkey tissues were determined by hybridization blotting analysis of total RNA from these tissues using a human apoAIV cDNA probe. This analysis showed that only fetal and adult intestine and adult but not fetal liver contain detectable amounts of apoAIV mRNA. These results indicate that the apoAIV gene evolved by amplification of an ancestral 66-bp sequence coding for a peptide with amphipathic properties and that conservation of the secondary structure and hydropathic properties of certain domains in apoAIV may be significant for the function(s) of this protein. Furthermore, these results indicate that in humans and nonhuman primates apoAIV mRNA synthesis occurs primarily in intestine while in liver apoAIV mRNA synthesis may be regulated by developmental and/or nutritional factors.

Apolipoprotein AIV (apoAIV) is a 46 000-dalton protein found initially in rat high-density lipoprotein (HDL) and chylomicrons (Swaney et al., 1974; Roheim et al., 1976) and later in human chylomicrons, very low density lipoproteins (VLDL), and in the $d > 1.21$ g/mL plasma fraction (Weisgraber et al., 1978; Beisiegel & Uterman, 1979; Green et al., 1979). Following ultracentrifugation, the majority of human apoAIV dissociates from lipoproteins and is found in the lipoprotein-free fraction of the plasma. In contrast, approximately 50% of rat and dog apoAIV remains associated with HDL under similar ultracentrifugation conditions (Roheim et al., 1976; Weisgraber et al., 1978). It has been reported that in rats the liver and intestine contribute 59% and 41%, respectively, in the plasma apoAIV pool (Wu & Windmueller, 1979). Intestinal synthesis of apoAIV has also been demonstrated in humans (Green et al., 1980). Newly synthesized intestinal apoAIV associates with chylomicrons and in humans

† This work was supported by grants from the National Institutes of Health (HL32032), the National Science Foundation (DCB-8400173), the March of Dimes Birth Defects Foundation (1-817), and the Massachusetts Affiliate of the American Heart Association (13-517-845). This research was performed at the Housman Medical Research Center of the Boston University Medical School and the Department of Cardiology at Children's Hospital. S.K.K. and V.I.Z. are Established Investigators of the American Heart Association. S.K.K. is a Syntex Scholar.

* Author to whom correspondence should be addressed.

0006-2960/86/0425-3962$01.50/0

accounts for 10–13% of the total chylomicron apolipoprotein (Green et al., 1980). Most of the rat chylomicron apoAIV is transferred to HDL upon entering circulation, whereas the bulk of human chylomicron apoAIV is released during catabolism of chylomicrons and is found in the lipoprotein-free fraction (Imaizumi et al., 1978; Green & Glickman, 1981). It has been shown that apoAIV is a better activator than apoAI of the enzyme lecithin:cholesterol acyltransferase (LCAT) when the acyl donor is L-phosphatidylcholine esterified with two saturated fatty acids (Steinmetz & Utermann, 1985). It is interesting that LCAT esterification of lipoprotein-bound cholesterol favors the retention of apoAIV on the lipoprotein particle whereas inhibition of LCAT activity is accompanied by dissociation of apoAIV from this particle (DeLamatre et al., 1983). Possible involvement of apolipoprotein CIII (apoCIII) in displacing apoAIV from lipoprotein particles has also been speculated (Weinberg & Spector, 1985).

The genes coding for apoAI, apoCIII, and apoAIV are closely linked and tandemly organized on the long arm of the human chromosome 11 (Karathanasis, 1985). In this paper a human apoAIV gene probe was used to isolate a nearly full-length apoAIV cDNA clone from an adult human liver library. In-frame translation of the cDNA sequence in this clone indicated that the human apoAIV consists of 396 amino acid residues including a 20 residue long signal peptide. The coding region of this cDNA sequence contains 15 nucleotide repeats, 11 of which code for 22 amino acid repeats with potentials of forming amphipathic helices. Although the rate of evolution of certain amino acid repeats in apoAIV is higher than the rate of evolution of the entire protein, the corresponding secondary structures and hydropathy profiles of these repeats are very similar between the human and rat apoAIV. A human apoAIV cDNA probe was used for hybridization blotting analysis of total RNA isolated from various human and monkey tissues. The results showed that apoAIV mRNA is present in fetal and adult intestine and in adult but not fetal liver while this mRNA is absent from adrenal, brain, spleen, ovaries, lung, kidney, heart, stomach, thymus, and muscle. These results indicate that the apoAIV gene evolved by amplification of an ancestral nucleotide sequence coding for a peptide with lipid binding properties and that evolutionary conservation of the higher order structures of certain domains in apoAIV may be significant for the function(s) of this protein. In addition, these results indicate that in humans and nonhuman primates apoAIV mRNA synthesis is specific for intestine while in liver apoAIV mRNA synthesis may be regulated by developmental and/or nutritional factors.

Experimental Procedures

Nucleotide Sequencing and Radioactive Labeling of DNA. DNA sequencing analysis was carried out either by the base-specific chemical cleavage method (Maxam & Gilbert, 1980) or by subcloning of various DNA fragments in M13 vectors (Messing et al., 1981) and using the resulting single-stranded DNA templates for DNA sequence determination by the dideoxy chain termination method (Sanger et al., 1977). Radioactive labeling of DNA was carried out by nick-translation (Rigby et al., 1977). Enzymes and radioactive material required for these analyses were purchased from New England Biolabs and New England Nuclear, respectively.

Tissues, Total RNA Preparation, and Hybridization Blotting Analysis. Various human fetal tissues were obtained from 20–22-week-old human abortuses, with no apparent malformations, under a protocol approved by the Research Advisory Committee of the Brigham and Women's Hospital, Boston, MA (Zannis et al., 1982). These tissues were collected within 1 h following the abortion and after freezing in liquid nitrogen were stored at −70 °C until further use. Adult human liver tissue was obtained from an organ transplant donor, victim of a car accident (a gift from Dr. Goldberger at Children's Hospital, Boston, MA). Cynomolgus monkey (*Macacca fascicularis*) tissues (a gift of Dr. K. C. Hayes, Brandeis University, Waltham, MA) were collected and stored in a similar fashion. HepG2 cells, SV40-transformed fibroblasts, U937 cells, and primary cultures of human monocyte macrophages were grown as described previously (Zannis et al., 1985). Cell pellets were collected and stored at −70 °C. Tissue or cell total RNA was prepared (Chirgwin et al., 1979), and its integrity was evaluated by the integrity of the 18S and 28S ribosomal RNAs upon electrophoresis in 1% agarose–formaldehyde gels (Goldberger, 1980) followed by staining with ethidium bromide. Following electrophoresis, the RNA was transferred onto nitrocellulose filters (Thomas, 1980) and hybridized with ^{32}P-labeled DNA probes as previously described (Zannis et al., 1985). The relative steady-state levels of apoAIV mRNA in various tissues were determined as previously detailed (Zannis et al., 1985) and expressed as percent of the apoAIV mRNA levels in the intestine.

Cell-Free Translation and Electrophoretic Analysis of ApoAIV. Human fetal intestine RNA was translated with a rabbit reticulocyte system (Bethesda Research Laboratories; Pelham & Jackson, 1976) and [^{35}S]methionine (1200 Ci/mmol, New England Nuclear). Immunoprecipitation of in vitro synthesized human apoAIV was carried out with a rabbit anti-human apoAIV antibody (provided generously by Dr. Karl Weisgraber) and a staphylococcal protein A preparation (Morris et al., 1982). The precipitant was mixed with 5 μg each of human and rat apoAIV (obtained from lymph chylomicrons of a single individual) and analyzed by two-dimensional gel electrophoresis (O'Farrell, 1975) followed by staining with Coomassie Brilliant Blue and autoradiography as described previously (Zannis et al., 1982). Determination of the isoelectric point and molecular weight of apoAIV was carried out as previously detailed (Zannis & Breslow, 1981).

Computer Analysis. Nucleotide sequence homologies were studied with the SEQ computer program (Butlag et al., 1982). Amino acid alignments were carried out with a previously described computer program (Wilbur & Lipman, 1983). Amino acid sequence homologies (Needleman & Wunsch, 1970), predictions of secondary structure (Chou & Fasman, 1982), and hydropathy profiles (Kyte & Doolittle, 1982) were carried out with computer programs developed by Intelligenetics.

Results

Isolation and Characterization of a Nearly Full-Length Human ApoAIV cDNA Clone. Screening of an adult human liver cDNA library with a probe containing human apoAIV gene sequences resulted in isolation of several clones, one of which (λgt10-2) was shown to contain a part of the human apoAIV cDNA sequence (Karathanasis, 1985). Another of these clones (λgt10-3) was purified and mapped with restriction enzymes (data not shown), and its cDNA sequence was determined (Figure 1A). In-frame translation of this sequence indicated that the clone λgt10-3 contains 51 nucleotides of the 5′-noncoding, 1188 nucleotides of the coding, and 164 nucleotides of the 3′-noncoding regions of the human apoAIV mRNA (Figure 1A). The amino acid sequence, derived from this clone, indicated that the human apoAIV contains 396 amino acid residues including a 20 residue long signal peptide (Figure 1A). Two-dimensional sodium dodecyl sulfate (SDS) gel electrophoretic comparison of in vitro syn-

A.

Amino Acid Residues		Nucleotide Residues	Nucleotide Repeats	Percent Homology
	GGG AGG ATC CAG TGT GGC AAG AAA CTC CTC CAG CCC AGC AAG CAG CTC AGG	51		
-20	ATG TTC CTG AAG GCC GTG GTC CTG ACC CTG GCC CTG GTG GCT GTC GCC GGA GCC AGG GCT MET Phe Leu Lys Ala Val Val Leu Thr Leu Ala Leu Val Ala Val Ala Gly Ala Arg Ala	111		
1	GAG GTC AGT GCT GAC CAG GTG GCC ACA GTG ATG TGG Glu Val Ser Ala Asp Gln Val Ala Thr Val MET Trp	147		
13	GAC TAC TTC AGC CAG CTG AGC AAC AAT GCC AAG GAG GCC GTG GAA CAT CTC CAG AAA Asp Tyr Phe Ser Gln Leu Ser Asn Asn Ala Lys Glu Ala Val Glu His Leu Gln Lys	204	I	59.6
32	TCT GAA CTC ACC CAG CAA CTC AAT GCC CTC TTC CAG GAC AAA CTT GGA GAA GTG AAC Ser Glu Leu Thr Gln Gln Leu Asn Ala Leu Phe Gln Asp Lys Leu Gly Glu Val Asn	261		40.3
51	ACT TAC GCA GGT GAC CTG CAG AAG AAG CTG GTG CCC TTT GCC ACC GAG CTG CAT GAA CGC CTG GCC Thr Tyr Ala Gly Asp Leu Gln Lys Lys Leu Val Pro Phe Ala Thr Glu Leu His Glu Arg Leu Ala	327	II	65.2
73	AAG GAC TCG GAG AAA CTG AAG GAG GAG ATT GGG AAG GAG CTG GAG GAG CTG AGG GCC CGG CTG CTG Lys Asp Ser Glu Lys Leu Lys Glu Glu Ile Gly Lys Glu Leu Glu Glu Leu Arg Ala Arg Leu Leu	393	III	69.7
95	CCC CAT GCC AAT GAG GTG AGC CAG AAG ATC GGG GAC AAC CTG CGA GAG CTT CAG CAG CGC CTG GAG Pro His Ala Asn Glu Val Ser Gln Lys Ile Gly Asp Asn Leu Arg Glu Leu Gln Gln Arg Leu Glu	459	IV	69.7
117	CCC TAC GCG GAC CAG CTG CGC ACC CAG GTC AAC ACG CAG GCC GAG CAG CTG CGG CGC CAG CTG GAC Pro Tyr Ala Asp Gln Leu Arg Thr Gln Val Asn Thr Gln Ala Glu Gln Leu Arg Arg Gln Leu Asp	525	V	72.7
139	CCC CTA GCA CAG CGC ATG GAG AGA GTG CTG CGG GAG AAC GCC GAC AGC CTG CAG GCC TCG CTG AGG Pro Leu Ala Gln Arg MET Glu Arg Val Leu Arg Glu Asn Ala Asp Ser Leu Gln Ala Ser Leu Arg	591	VI	57.6
161	CCC CAC GCC GAC GAG CTC AAG GCC AAG ATC GAC CAG AAC GTG GAG GAG CTC AAG GGA CGC CTT ACG Pro His Ala Asp Glu Leu Lys Ala Lys Ile Asp Gln Asn Val Glu Glu Leu Lys Gly Arg Leu Thr	657	VII	81.8
183	CCC TAC GCT GAC GAA TTC AAA GTC AAG ATT GAC CAG ACC GTG GAG GAG CTG CGC CGC AGC CTG GCT Pro Tyr Ala Asp Glu Phe Lys Val Lys Ile Asp Gln Thr Val Glu Glu Leu Arg Arg Ser Leu Ala	723	VIII	77.3
205	CCC TAT GCT CAG GAC ACG CAG GAG AAG CTC AAC CAC CAG CTT GAG GGC CTG ACC TTC CAG ATG AAG Pro Tyr Ala Gln Asp Thr Gln Glu Lys Leu Asn His Gln Leu Glu Gly Leu Thr Phe Gln MET Lys	789	IX	69.7
227	AAG AAC GCC GAG GAG CTC AAG GCC AGG ATC TCG GCC AGT GCC GAG GAG CTG CGG CAG AGG CTG GCG Lys Asn Ala Glu Glu Leu Lys Ala Arg Ile Ser Ala Ser Ala Glu Glu Leu Arg Gln Arg Leu Ala	855	X	72.7
249	CCC TTG GCC GAG GAC GTG CGT GGC AAC CTG AAG GGC AAC ACC GAG GGG CTG CAG AAG TCA CTG GCA Pro Leu Ala Glu Asp Val Arg Gly Asn Leu Lys Gly Asn Thr Glu Gly Leu Gln Lys Ser Leu Ala	921	XI	69.7
271	GAG CTG GGT GGG CAC CTG GAC CAG CAG GTG GAG GAG TTC CGA CGC CGG GTG GAG Glu Leu Gly Gly His Leu Asp Gln Gln Val Glu Glu Phe Arg Arg Arg Val Glu	975	XII	77.8
289	CCC TAC GGG GAA AAC TTC AAC AAA GCC CTG GTG CAG CAG ATG GAA CAG CTC AGG CAG AAA CTG GGC Pro Tyr Gly Glu Asn Phe Asn Lys Ala Leu Val Gln Gln MET Glu Gln Leu Arg Gln Lys Leu Gly	1041	XIII	65.2
311	CCC CAT GCG GGG GAC GTG GAA GGC CAC TTG AGC TTC CTG GAG AAG GAC CTG AGG GAC AAG GTC AAC Pro His Ala Gly Asp Val Glu Gly His Leu Ser Phe Leu Glu Lys Asp Leu Arg Asp Lys Val Asn	1107	XIV	60.6
333	TCC TTC TTC AGC ACC TTC AAG GAG AAA GAG AGC CAG GAC AAG ACT CTC TCC CTC Ser Phe Phe Ser Thr Phe Lys Glu Lys Glu Ser Gln Asp Lys Thr Leu Ser Leu	1161		42.6
351	CCT GAG CTG GAG CAA CAG CAG GAA CAG CAG CAG GAG CAG CAG CAG GAG CAG GTG CAG ATG CTG GCC Pro Glu Leu Glu Gln Gln Gln Glu Gln Gln Gln Glu Gln Gln Gln Glu Gln Val Gln MET Leu Ala	1227	XV	65.2
373	CCT TTG GAG AGC TGA GCT GCC CCT GGT GCA CTG GCC CCA CCC TCG TGG ACA CCT GCC CTG CCC Pro Leu Glu Ser END	1290		
	TGC CAC CTG TCT GTC TGT CCC AAA GAA GTT CTG GTA TGA ACT TGA GGA CAC ATG TCC AGT GGG	1353		
	AGG TGA GAC CAC CTC TCA ATA TTC AAT AAA GCT GCT GAG AAT CTA GCC TCA AAA AAA AAA AAA	1416		

B. CONSENSUS SEQUENCE

```
                /   //    /  /                       / /       /
      G     C   C   GG    A C A C G C C C          A  G G A      AC
CCC TAG GCC GAG GAG CTG AAG GAC AAG CTG GAG CAG AAG GTG GAG GAG CTG CGG CAC CGG CTG GCG
Pro Tyr Ala Glu Glu Leu Lys Asp Lys Leu Glu Glu Lys Val Glu Glu Leu Arg His Arg Leu Glu
                Asp                     Asp Gln Gln                 Gln Asp         Ala
                                                                        Gln
            -   -   L   +   -   +   L               -   -   L   +   ±   +   L
|______________________________________||_________________________________________|
                  α                                         β
```

FIGURE 1: (A) Alignment of the clone λgt10-3 cDNA sequence in segments (nucleotide repeats I–XV) and their homology (percent homology) to a consensus sequence generated (see panel B) with the nucleotide sequences of these segments. An amino acid sequence (three-letter code) derived by in-frame translation of the λgt10-3 cDNA sequence is shown below the nucleotide sequence. Nucleotide and amino acid residues are numbered along the right and left sides of these sequences, respectively. Signal peptide sequences are indicated by negative amino acid residue numbers. (B) A consensus sequence derived from nucleotide repeats present in the λgt10-3 cDNA. Single residues in the consensus sequence occur with a frequency of ≥0.5 among corresponding residues in these repeats. Presence of two residues at the same position of this consensus sequence indicates that the bottom residues occur with frequencies of 0.4–0.5 while the top residues occur with frequencies of 0.33 (primed residues) or 0.4 (nonprimed residues). All residues at nucleotide position 9 and 65 occur with a frequency of 0.36 while at position 23 the primed residue occurs with a frequency of 0.27. An amino acid sequence derived by in-frame translation of this nucleotide consensus sequence (including alternative nucleotide residues) is also shown. This sequence is subdivided in 11-residue segments (α and β) to show the symmetric preservation of their charged and hydrophobic residues.

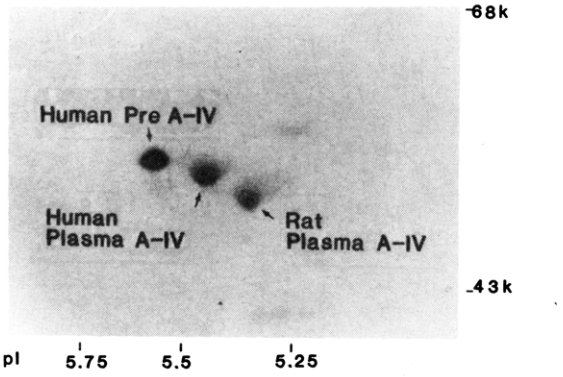

FIGURE 2: Two-dimensional polyacrylamide gel electrophoresis and autoradiography of proteins immunoprecipitated from the translation cocktail with specific anti-apoAIV antibodies. A 50-μL aliquot of the translation cocktail of fetal intestinal mRNA was immunoprecipitated with rabbit anti-human apoAIV. The apoAIV antibody complex was extracted with lysis buffer, mixed with 5 μg each of human and rat chylomicron apoAIV, and analyzed by two-dimensional polyacrylamide gel electrophoresis following autoradiography as described under Experimental Procedures. The slab gel obtained from this analysis was stained with Coomassie Brilliant Blue to identify the plasma forms of human and rat apoAIV and then autoradiographed to identify the cell-free translation product of human apoAIV mRNA. The autoradiogram obtained from this analysis superimposed on the corresponding slab gel is shown. This superimposition shows the molecular weight and charge relationship of the plasma and cell-free translated apoAIV. Molecular weight markers are bovine serum albumin (68K) and ovalbumin (43K). The isoelectric points were determined as previously described (Zannis et al., 1981).

thesized (by translation of human fetal intestine total RNA) and plasma forms of the human apoAIV shows that the primary translation product of the human apoAIV mRNA has a higher molecular weight and differs by a 2+ charge compared to its mature plasma counterpart (Figure 2). These results are consistent with cotranslational cleavage of the 20 residue signal peptide of the human pre-apoAIV (Figure 1A).

Analysis of the λgt10-3 cDNA sequence for internal homologies indicated the presence of 15 (I–XV) nucleotide segments with significant ($E < 0.001$) homology to a nucleotide consensus sequence derived from these segments (Figure 1). Significant ($E < 0.001$) homology of segments VI and XIV was obtained by using the alternative nucleotides in the consensus sequence (Figure 1). This analysis also indicated that the nucleotide segments I and XII are truncated at their 3′ and 5′ ends, respectively (Figure 1A). Similar comparisons of the cDNA-derived human apoAIV amino acid sequence, with the program ALIGN (Intelligenetics), indicated that the nucleotide segments I, II, V, VII, VIII, IX, X, XII, XIII, XIV, and XV code for amino acid segments with extensive homology to an amino acid consensus sequence generated with these amino acid segments (data not shown). A 22 amino acid long sequence was derived from the consensus sequence of the nucleotide repeats in λgt10-3 cDNA (Figure 1B). This amino acid sequence consists of two symmetrical 11 residue long segments (α and β) with similar arrangements of charged and hydrophobic amino acid moieties. In addition, analysis of the secondary structure and hydropathy profile of this sequence indicated that it has the potential of nucleating in an α-helical structure with alternating hydrophobic and hydrophilic residues [Karathanasis (1985) and data not shown]. This arrangement is typical of amphipathic structures, and it has been suggested that such structures may be involved in lipid binding (Kaiser & Kezdy, 1983).

These results suggest that the apoAIV gene evolved by intragenic amplification of an ancestral 66-bp segment coding for a peptide with lipid binding properties and that this 66-bp DNA segment may have itself evolved by tandem duplication of a primordial 33 bp long DNA sequence.

Comparison between Human and Rat ApoAIV Amino Acid Sequences. Differences between amino acid sequences related by descent (orthologous sequences) reflect the phylogenetic branching order of the species in which these sequences are found and may underline differences of the function(s) of the corresponding proteins. Alignment of the human and rat (Boguski et al., 1984) apoAIV amino acid sequences, using a previously described computer method (Wilbur & Lipman, 1983), indicated a five-residue deletion (residues 388–392) near the carboxy terminus of the rat sequence (Figure 3). This difference is consistent with the different electrophoretic mobilities of the mature plasma forms of the human and rat apoAIV in two-dimensional SDS–polyacrylamide gels (Figure 2). This alignment also indicated that these proteins are 61.8% homologous (Figure 3). The 38.2% amino acid difference between these proteins corresponds to a PAM value of 52 (Dayhoff, 1978). Since the rodent and primates diverged from their common ancestor approximately 80 million years ago (Romer, 1966), it can be calculated that the rate of evolution of apoAIV is 65 PAMs per 100 residues per 100 million years.

Alignment of the human and rat apoAIV amino acid sequences according to the nucleotide repeats in the λgt10-3 cDNA allowed estimation of the rates of evolution of various repeated and nonrepeated domains in this protein (Figure 3). The results show that the rates of evolution of repeats III, V, VI, VIII, and X are higher than the rate of evolution of the entire protein and that the rates of evolution of repeats V, VI, and X differ by 3–5 standard deviations from the average rate of evolution of the other repeats in apoAIV. To determine whether the high rate of evolution of these repeats corresponds to differences in the higher order structures of human and rat apoAIV, the secondary structure and hydropathy profiles of these proteins were generated, aligned, and visually compared (Figure 4). The results show that the secondary structures and hydropathy profiles of the repeats III, IV, V, X, and XIV but not VI and VIII are very similar between the human and rat apoAIV (Figure 4). It is interesting that the repeats VI and VIII flank the repeat VII, a potential LCAT activator in the rat apoAIV (Boguski et al., 1984). It should also be pointed out that although the rates of evolution of repeats VII, XII, and XV are similar to or lower than that for the entire apoAIV, the secondary structures and hydropathy profiles of these repeats appear to be different between the human and rat apoAIV (Figure 4). These results indicate that conservation of the higher order structures of certain domains together with conservation of the primary structures of other domains in apoAIV may be significant for the function(s) of this protein.

Tissue-Specific Expression of ApoAIV Gene. Differences in the relative steady-state level of an mRNA species in different organs are related to both the different controls involved in the tissue-specific expression of the corresponding gene and the role of the encoded protein in the overall physiology of the organism. Total RNA was isolated from various fetal human tissues blotted and hybridized with human apoAIV or apoCIII (Karathanasis et al., 1985) cDNA probes. The results indicate that apoAIV mRNA is present in fetal intestine but not in liver, adrenal, brain, spleen, ovaries, lung, kidney, heart, stomach, thymus, or muscle (Figure 5A). The presence of apoCIII mRNA in fetal human liver is compatible with previous results (Zannis et al., 1985) and contrasts the absence of apoAIV mRNA in this tissue. Similarly, hybridization blotting analysis of total RNA from various adult human and

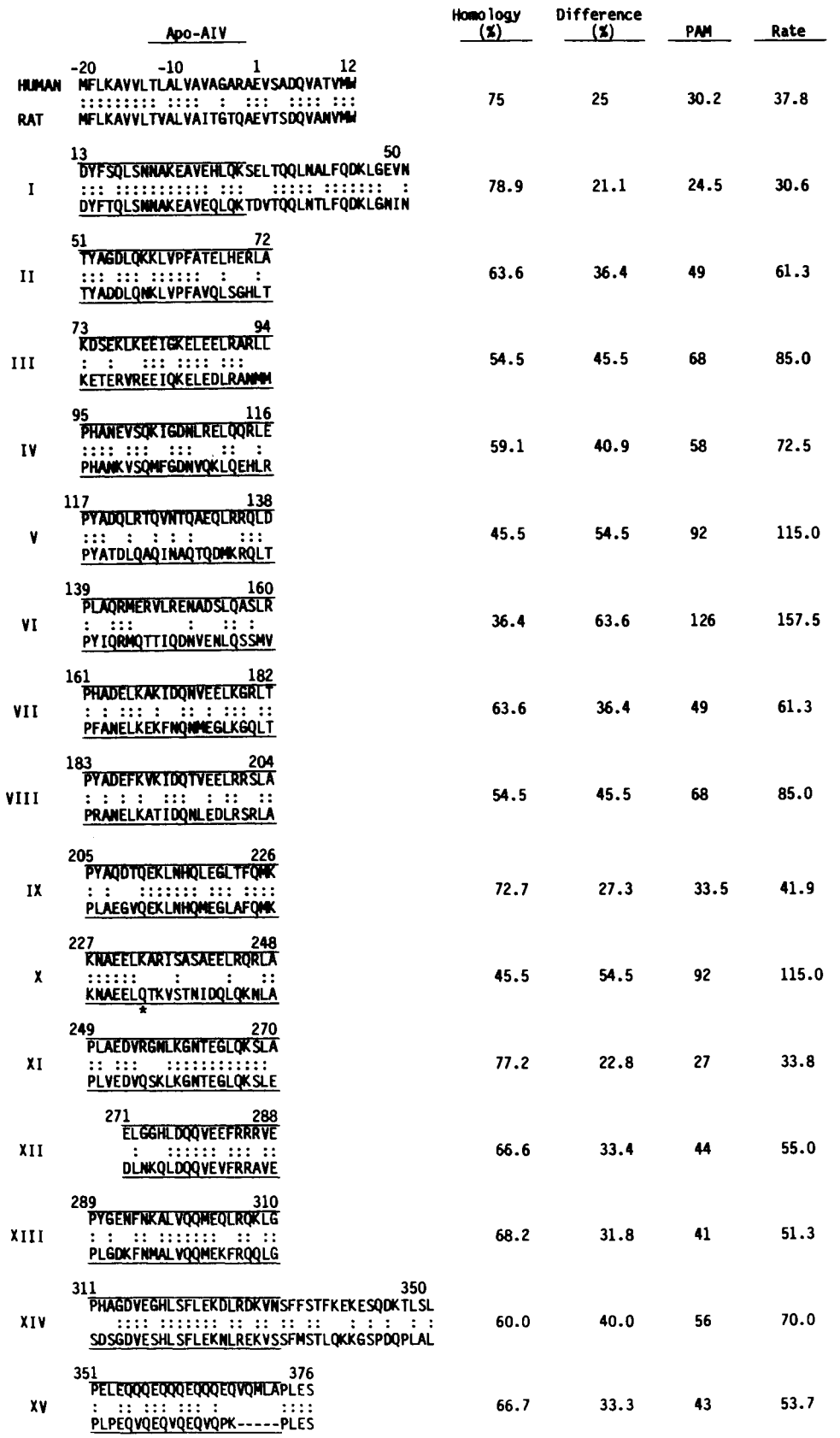

FIGURE 3: Alignment of the human and rat apoAIV amino acid sequences (single-letter code). This alignment was obtained by a previously described computer method (Wilbur & Lipman, 1983). The parameters used for this alignment are Ktuple = 1, window = 20, and gap penalty = 7. Identical residues in these sequences are denoted by colons. A difference between the cDNA- (Bogusky et al., 1984) and gene- (data not shown) derived rat apoAIV amino acid sequences is indicated by an asterisk. The amino acid segments corresponding to the nucleotide repeats (I–XV) in the human and rat apoAIV cDNA are overlined and underlined, respectively. Negative and positive numbers above the sequences indicate the signal peptide and mature apoAIV amino acid sequences, respectively. The percent residue homology, percent residue difference, the corresponding PAM value (Dayhoff, 1978), and the rate of evolution (assuming that rodents and primates separated 80 million years ago; Romer, 1966) of the aligned repeats of the human and rat apoAIV are shown.

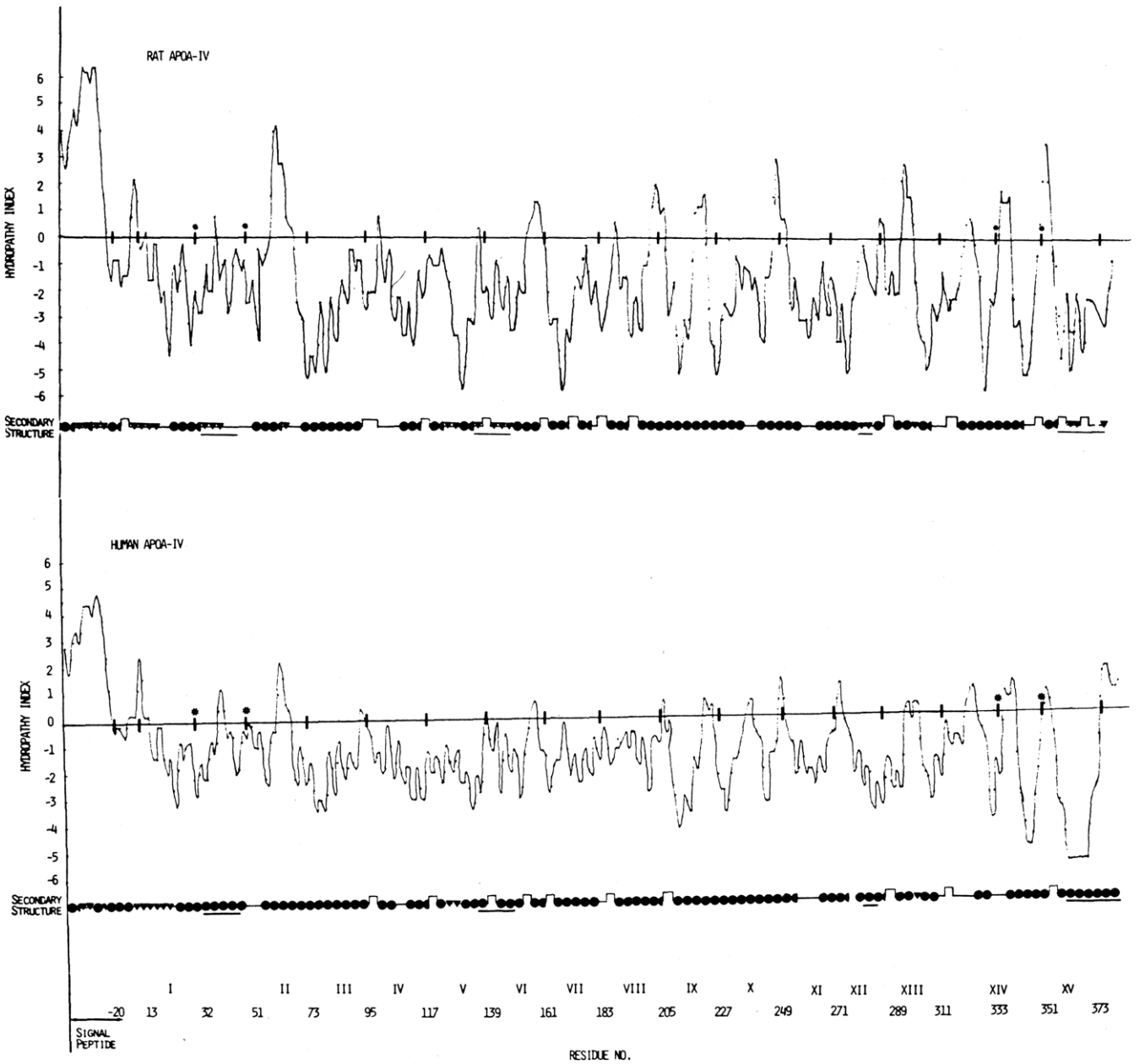

FIGURE 4: Comparison of the predicted human and rat apoAIV hydropathy profiles and secondary structures. Hydropathy profiles (Kyte & Doolittle, 1982) and secondary structures (Chou & Fasman, 1978) were generated as described under Experimental Procedures. Roman numerals show the amino acid segments coded by the nucleotide repeats in the rat and human apoAIV cDNA sequences. Secondary structure is shown as follows: (●) three residues of α-helix; (▼) three residues of β-structure; (⊓) β-turn or loop structure; (—) random coil. Differences between the secondary structure of these proteins are underlined.

monkey tissues was carried out with a human apoAIV cDNA probe. The results indicate that apoAIV mRNA is present primarily in monkey intestine (adult human intestine also contains apoAIV mRNA; data not shown) and to a lesser extent in liver but not in monocyte macrophage, muscle, kidney, lung, or adrenal (Figure 5B). Furthermore, none of the various established cell lines examined showed detectable amounts of apoAIV mRNA (Figure 5). The size of human or monkey apoAIV mRNA is approximately 1700 nucleotides long (Figure 5B). Since the cDNA sequence of the clone λgt10-3 is approximately 1500 bp long, it can be estimated that the poly(A) tail of the human or monkey apoAIV mRNA is approximately 200 nucleotides long. Quantitation (Zannis et al., 1985) of the relative steady-state apoAIV mRNA levels in various human tissues indicated that the adult liver contains 5% of the apoAIV mRNA levels found in intestine. These results indicate that in human and nonhuman primates apoAIV mRNA synthesis is specific for intestine, while in liver expression of the apoAIV gene may be regulated by developmental and/or nutritional factors.

Discussion

The genes coding for the apolipoproteins AI, CIII, and AIV are closely linked and tandemly organized on the long arm of human chromosome 11 (Karathanasis, 1985; Bruns et al., 1984; Cheung et al., 1984; Law et al., 1984). In this study, a probe containing human apoAIV gene sequences was used to isolate a nearly full-length apoAIV cDNA clone from an adult human liver library. In-frame translation in the cDNA sequence of this clone indicated that the human apoAIV contains 396 amino acid residues including a 20-residue signal peptide. The signal peptide and a portion of the mature human apoAIV amino acid sequence has been reported previously (Gordon et al., 1984). However, at residue positions −4, 3,

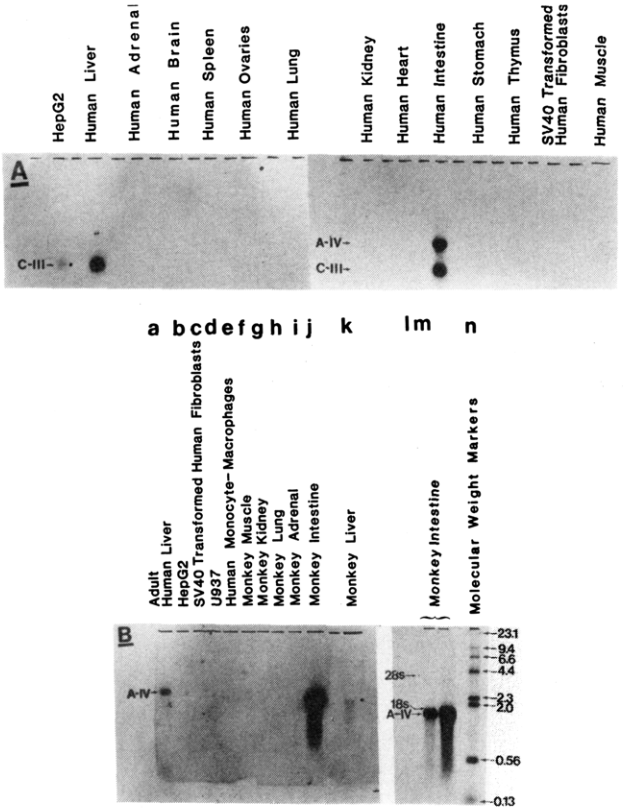

FIGURE 5: Blotting analysis of total RNA isolated from various human and monkey tissues as well as various cell lines of human origin. (A) Total RNA (15 μg/lane) isolated from various fetal human tissues (shown) and cell lines of human origin (HepG2, SV40-transformed human fibroblasts) was electrophoresed in agarose–formaldehyde gels, blotted, and hybridized with a human apoAIV (clone λgt10-3) or apoCIII (Karathanasis et al., 1985) cDNA probe. The resulting autoradiograms (superimposed on each other) are shown. (B) Total RNA isolated from various adult human or monkey tissues (shown) and cell lines of human origin (HepG2, SV40-transformed human fibroblasts and U937) was electrophoresed in agarose–formaldehyde gels, blotted, and hybridized with a human apoAIV cDNA probe (see panel A). The resulting autoradiogram is shown. To show the relative abundance of apoAIV mRNA in different tissues, different amounts of total RNA were used in each lane of these RNA blots. Specifically, lanes a and b contain 15 μg, lanes c–k contain 30 μg, and lanes l–m contain 2 and 6 μg of total RNA, respectively. Molecular weight markers (lane n) are 5′ end labeled *Hin*dIII-digested λDNA (New England Biolabs).

4, and 9 the previously reported sequence (Gordon et al., 1984) contains the residues Leu, Ile, Gly, and Asn while the cDNA-derived sequence (Figure 1A) indicates Gly, Ser, Ala, and Thr, respectively. It is not clear whether these differences are due to amino acid polymorphisms in human apoAIV.

It has been noted previously that different apoAIV cDNA clones isolated from the same human liver library contain either the nucleotide T or C at position 1186 (Figure 1A) and that the presence of T in this position generates a termination codon resulting in an 18 residue shorter version of the human apoAIV (Karathanasis, 1985). The nucleotide sequence of the apoAIV cDNA clone λgt10-3 contains the nucleotide C at position 1186 and thus codes for the longer version of human apoAIV (Figure 1A). It is interesting that for several different cDNA and genomic apoAIV clones the presence of the nucleotide T or C at position 1186 is linked to the presence of the sequence $(CTGT)_4$ or $(CTGT)_3$, respectively, between nucleotide positions 1296 and 1309 in the 3′ untranslated region of the apoAIV mRNA (Figure 1A and unpublished results). However, the presence, linkage, and frequency of these DNA polymorphisms as well as their correspondence to short or long versions of the human apoAIV have not been studied in different human populations.

It has been previously reported that the human and rat apoAI (Karathanasis et al., 1983; Boguski et al., 1985), apoE (Das et al., 1985; Paik et al., 1985), and apoAIV (Boguski et al., 1984; Karathanasis, 1985) genes contain 15 nucleotide repeats with consensus sequences highly homologous to each other (Karathanasis, 1985). Computer search for internal nucleotide homolgies in the human (Figure 1A) or rat (Boguski et al., 1984) apoAIV cDNA sequences indicated that the coding region of the human or rat (data not shown) apoAIV cDNA contains 15 (I–XV) nucleotide repeats, 11 of which code for 22 amino acid repeats in the corresponding proteins. Similar search for internal amino acid homologies also indicated the presence of amino acid repeats in the rat apoAIV (Boguski et al., 1984). However, the amino acid based analysis did not detect the repeats XIV and XV and indicated different arrangements for the repeats I, II, III, and XII. It is therefore possible that these discrepancies may indicate differences of the rates of silent and/or amino acid replacement nucleotide substitutions among the various repeats in the human and rat apoAIV genes.

An amino acid sequence derived from the consensus sequence of the repeated nucleotide segments in the human apoAIV cDNA has similar features with amino acid consensus sequences generated from the human apoAI (Karathanasis et al., 1983) and apoE (Das et al., 1985) and rat apoAIV (Boguski et al., 1984). More specifically, these amino acid sequences consist of two symmetrical 11 residue long segments with similar arrangement of charged and hydrophobic residues (α and β in Figure 1B). In addition, the secondary structure and hydropathy profile of this amino acid sequence indicated that it can form amphipathic helices (Karathanasis, 1985) possibly involved in lipid binding (Kaiser & Kezdy, 1983). These results suggest that the apoAIV gene evolved by intragenic amplification of an ancestral 66-bp DNA segment coding for a peptide with lipid binding properties and that this segment may have, itself, evolved by tandem duplication of a primordial 33 bp long DNA sequence.

A fundamental difference between the rat and the human apoAIV is their different affinity for the HDL particles present in the plasma of these species. Thus, following ultracentrifugation 50% of rat apoAIV is found in HDL whereas the majority of human apoAIV is recovered in the $d > 1.21$ g/mL fraction (Roheim et al., 1976; Weisgraber et al., 1978). Computer-aided alignment of the human and rat apoAIV amino acid sequences showed a five amino acid deletion close to the carboxy terminus (amino acids 388–392) of the rat apoAIV. This size difference between human and rat apoAIV is compatible with the different electrophoretic mobilities of the two proteins in two-dimensional polyacrylamide gels. The cDNA-derived amino acid sequences of the two proteins indicate that the overall charge of the plasma form of human and rat apoAIV is 9– and 8–, respectively. On the basis of the net charge value alone, one would predict that the human apoAIV has a more acidic isoelectric point, and thus, it ought to focus to the right and not to the left of the rat apoAIV (Figure 2). The reason for the observed discrepancy is not clear. It may be due to amino acid polymorphisms in these proteins; for instance, sequencing of the rat apoAIV gene (data not shown) indicated Gln at amino acid residue position 233 while the previously reported sequence (Boguski et al., 1984) indicated His at the same location. Alternatively, this abnormal isoelectric focusing behavior could result from subtle conformational differences that may result in differential

shielding of the charged residues of these proteins. An example of similarly abnormal isoelectric focusing behavior can be found in human apoAII and apoCII (Zannis & Breslow, 1984). Although both these proteins have similar size and a total of four negative charges, the apoAII has a more basic apparent isoelectric point than the apoCII.

Alignment and comparison of the human and rat cDNA-derived amino acid sequences indicated that these proteins are 61.8% homologous. The 38.2% difference between these proteins corresponds to a PAM value of 52 (Dayhoff, 1978) and a rate of evolution of 65 PAMs per 100 residues per 100 million years (assuming that the rodents and primates separated 80 million years ago; Romer, 1966). Similar alignment and comparison of the human and rat apoAI amino acid sequences indicated that the rate of evolution of this protein is 60.5 PAMs per 100 residues per 100 million years (unpublished results). It therefore appears that apoAIV and apoAI evolved with similar rates since the primate–rodent split. However, comparison of the human and dog (Chung et al., 1981) apoAI sequences indicated a rate of evolution at 23 PAMs per 100 residue per 100 million years (unpublished results). An even lower rate of evolution for apoAI (16 PAMs per 100 residues per 100 million years) has been previously calculated on the basis of short apoAI amino acid segments from different species (Barker & Dayhoff, 1977). Since the carnivora branched off before or around the primate–rodent split (Romer, 1966), it can be calculated that the rate of amino acid substitutions in apoAIV and apoAI is 2.8 (65/23) and 2.6 (60.5/23) times faster, respectively, in the rat than in the human. A similar conclusion was recently reached by comparing the rates of synonymous nucleotide substitutions in 11 different genes, and it was suggested that the shorter generation times of rodents compared to primates may be responsible for the higher rates of evolution of the rodent genes (Wu & Li, 1985).

The rates of evolution of certain amino acid repeats (III, V, VI, VIII, and X) in apoAIV are higher than the rate of evolution of the entire protein (Figure 3). However, the corresponding secondary structures and hydropathy profiles of most of these repeats (III, V, and X) are very similar between the human and rat proteins. It therefore appears that conservation of the higher order structure of these repeats may be more important than conservation of their primary structure for the function(s) of this protein. In contrast, although the rate of evolution of other repeats (VII, XII, and XV) in apoAIV is similar to or lower than the rate of evolution of the entire protein, it appears that the corresponding secondary structure and/or hydropathy profiles of these repeats are different between the human and rat proteins. It may therefore be speculated that conservation of the primary structures of these repeats may be more important than conservation of their higher order structures for the function(s) of this protein. Thus, the conservation of the primary structure of repeat VII and its possible involvement in LCAT activation (Boguski et al., 1984) may indicate that primary rather than higher order protein structures may be involved in LCAT activation. It, however, remains to be established whether these structural differences between the human and rat apoAIV are also related to possible differences in the function(s) of these proteins.

It has been previously shown that in rat the liver and intestine contribute 59% and 41%, respectively, to the apoAIV plasma pool (Wu & Windmueller, 1979). However, blotting analysis of total RNA from various human and monkey tissues, using a human apoAIV cDNA probe, indicated that adult liver contains 5% of the apoAIV mRNA levels present in intestine. Similar results have also been obtained by quantitative dot blot hybridization analysis of adult rat liver and intestine total RNA with a rat apoAIV cDNA probe (unpublished results). It is likely that the differences between apoAIV protein and mRNA determinations in liver and intestine reflect the size differences of these organs in humans or rats.

It is finally interesting to point out that although the genes coding for apoAI, apoCIII, and apoAIV are closely linked in the human genome (Karathanasis, 1985), it appears that their expression is regulated differently. For instance, previous studies (Zannis et al., 1985) showed that the apoAI gene is expressed in a variety of human tissues such as fetal, adrenal, kidney, and heart, which do not express the apoAIV and apoCIII genes. In addition, both apoAI and apoCIII genes are expressed at high levels in fetal liver whereas apoAIV is not. Clearly, more work is needed to elucidate putative developmental, nutritional, and other factors involved in regulation of expression of apoAI, apoCIII, and apoAIV genes.

ACKNOWLEDGMENTS

We thank Elizabeth Walsh and Tracy Fitzpatrick for their expert assistance, Marie Fennell for typing the manuscript, and Emily Flynn for the art work.

Registry No. DNA (human apolipoprotein AIV messenger RNA complementary), 102305-79-3; apolipoprotein AIV (human protein moiety precursor), 102305-81-7; apolipoprotein AIV (human protein moiety), 102305-80-6.

REFERENCES

Barker, W. C., & Dayhoff, M. O. (1977) *Comp. Biochem. Physiol., B: Comp. Biochem. B57*, 309–315.

Beisiegel, U., & Uterman, G. (1979) *Eur. J. Biochem. 93*, 601.

Boguski, M. S., Elshourbagy, N., Taylor, J. M., & Gordon, J. I. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5021–5025.

Boguski, M. S., Elshourbagy, N., Taylor, J. M., & Gordon, J. I. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 992–996.

Bruns, G. A. P., Karathanasis, S. K., & Breslow, J. L. (1984) *Arteriosclerosis 4*, 97–102.

Brutlag, D. L., Clayton, S., Friedland, P., & Kedes, L. H. (1982) *Nucleic Acids Res. 10*, 279–294.

Cheung, P., Kao, F.-T., Law, M. L., Jones, C., Puck, T. T., & Chan, L. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 508–511.

Chirgwin, J. M., Przybla, A. E., MacDonald, R. Y., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–148.

Chung, H., Randolph, A., Reardon, I., & Heinrikson, R. L. (1982) *J. Biol. Chem. 257*, 2961–2967.

Das, H. K., McPherson, J., Bruns, G. A. P., Karathanasis, S. K., & Breslow, J. L. (1985) *J. Biol. Chem. 260*, 6240–6247.

Dayhoff, M. O., Eck, R. V., & Park, C. M. (1978) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, Suppl. 3, pp 75–84, Georgetown University Medical Center, Washington, DC.

DeLamatre, J. G., Hoffmeier, C. A., Lacko, A. G., & Roheim, P. S. (1983) *J. Lipid Res. 24*, 1578–1585.

Dvorin, E., Mantulin, W. M., Rohde, M. F., Gotto, A. M., Jr., Pownall, H. J., & Sherrill, B. C. (1985) *J. Lipid Res. 26*, 38–46.

Goldberg, D. A. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5794–5798.

Gordon, J. I., Bisagaier, C. L., Sims, H. F., Sachdev, O. P., Glickman, R. M., & Strauss, A. W. (1984) *J. Biol. Chem. 259*, 468–474.

Green, P. H. R., & Glickman, R. M. (1981) *J. Lipid Res. 22*, 1153–1173.

Green, P. H. R., Glickman, R. M., Saudek, C. D., Blum, C. B., & Tall, A. R. (1979) *J. Clin. Invest. 64*, 233–242.

Green, P. H. R., Glickman, R. M., Riley, J. W., & Quinet, E. (1980) *J. Clin. Invest. 65*, 911–919.

Imaizumi, K., Fainaur, M., & Havel, R. J. (1978) *J. Lipid Res. 19*, 712–722.

Kaiser, E. T., & Kezdy, F. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1137–1143.

Karathanasis, S. K. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6374–6378.

Karathanasis, S. K., Zannis, V. I., & Breslow, J. L. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6147–6151.

Karathanasis, S. K., Zannis, V. I., & Breslow, J. L. (1985) *J. Lipid Res. 26*, 451–456.

Kyte, S., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.

Law, S. W., Gray, G., Brewer, H. B., Jr., Sakaguchi, A. Y., & Naylor, S. L. (1984) *Biochem. Biophys. Res. Commun. 118*, 934–942.

Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Messing, J., Crea, K., & Seeburg, H. P. (1981) *Nucleic Acids Res. 9*, 309–319.

Morris, K. M., Aden, D. P., Knowles, B. B., & Colten, H. R. (1982) *J. Clin. Invest. 70*, 906–913.

Needleman, S. B., & Wunsch, C. C. (1970) *J. Mol. Biol. 48*, 443–453.

O'Farrell, P. H. (1975) *J. Biol. Chem. 250*, 4007–4021.

Paik, Y.-K., Chang, D. J., Reardon, C. A., Davies, G. E., Mahley, R. W., & Taylor, J. M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3445–3449.

Pelham, H. R. B., & Jackson, R. J. (1976) *Eur. J. Biochem. 67*, 247–256.

Rigby, P. W. J., Dieckman, M., Rhodes, C., & Berg, P. (1977) *J. Mol. Biol. 113*, 237–251.

Roheim, P. S., Edelstein, D., & Pinter, G. G. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 1757–1760.

Romer, A. S. (1966) *Vertebrate Paleontology*, University of Chicago Press, Chicago.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Steinmetz, A., & Utermann, G. (1985) *J. Biol. Chem. 260*, 2258–2264.

Swaney, J. B., Reese, H., & Eder, H. A. (1974) *Biochem. Biophys. Res. Commun. 59*, 513–519.

Thomas, P. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5201–5205.

Weinberg, R. B., & Spector, M. S. (1985) *J. Lipid Res. 26*, 26–37.

Weisgraber, K. H., Bersot, T. P., & Mahley, R. W. (1978) *Biochem. Biophys. Res. Commun. 85*, 287.

Wilbur, W. J., & Lipman, D. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 726–730.

Wu, A. L., & Windmueller, H. G. (1979) *J. Biol. Chem. 254*, 7316–7322.

Wu, C.-I., & Li, W.-H. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1741–1745.

Zannis, V. I., & Breslow, J. L. (1981) *Biochemistry 20*, 1033–1041.

Zannis, V. I., & Breslow, J. L. (1984) *Adv. Hum. Genet. 14*, 125–215.

Zannis, V. I., Kurnit, D., & Breslow, J. L. (1982) *J. Biol. Chem. 257*, 536–544.

Zannis, V. I., Cole, F. S., Jackson, C. L., Kurnit, D. M., & Karathanasis, S. K. (1985) *Biochemistry 24*, 4450–4455.