Isolation and characterization of cDNA clones corresponding to two different human apoC-III alleles

Sotirios K. Karathanasis,*, Vassilis I. Zannis,*,**,††and Jan L. Breslow*,**,†**

Metabolism Division, Children's Hospital,* Harvard Medical School,** Boston, MA 02115; Laboratory of Biochemical Genetics and Metabolism, Rockefeller University,† New York, NY 10021; and Section of Molecular Genetics, Cardiovascular Institute, Boston University Medical Center,†† 80 East Concord St., Boston, MA 02118

Abstract We have recently reported that the human apolipoprotein A-I (apoA-I) and apolipoprotein C-III (apoC-III) genes are physically linked and that the presence of a DNA insertion in the apoA-I gene is correlated with apoA-I–apoC-III deficiency in patients with premature atherosclerosis. In addition, the presence of a polymorphic restriction endonuclease site (SacI) in the 3′ noncoding region of apoC-III mRNA has been correlated with hypertriglyceridemia in humans. In this study, we report the isolation and characterization of cDNA clones containing the entire apoC-III mRNA coding sequence. The nucleotide-derived apoC-III amino acid sequence indicates that the apoC-III primary translational product contains a 20 amino acid N-terminal extension, which conforms with the general properties of known signal peptides, and is highly homologous to the recently reported rat apoC-III signal peptide. The DNA-derived apoC-III amino acid sequence differs from the previously reported apoC-III amino acid sequence at four amino acid residues. More specifically, at positions +32, +33, +37, +39, the DNA sequence predicts Glu, Ser, Gln, Ala, respectively, while the previously reported sequence specifies Ser, Gln, Ala, Gln, respectively. Finally, isolation and characterization of apoC-III cDNA clones, with or without the polymorphic SacI restriction site, indicated that the apoC-III nucleotide sequence corresponding to the Sac^+ and Sac^- clones differs at three nucleotide sites; however, the amino acid sequence specified by the Sac^+ and Sac^- alleles is identical. — **Karathanasis, S. K., V. I. Zannis and J. L. Breslow.** Isolation and characterization of cDNA clones corresponding to two different human apoC-III alleles. *J. Lipid Res.* 1985. **26:** 451–456.

Supplementary key words apoC-III mRNA • nucleotide-derived amino acid sequence • hypertriglyceridemia

ApoC-III is a 79 amino acid glycoprotein of known primary structure (1, 2). The carbohydrate to apoC-III molar ratio is 1 for galactose, 1 for galactosamine, and 0, 1, or 2 for sialic acid. These variations in sialic acid apoC-III molar ratios result in three apoC-III isoforms designated CIII-0, CIII-1, and CIII-2, respectively (3, 4). The importance of apoC-III sialation remains unclear, although an earlier report indicated that type V hyperlipoproteinemia is associated with an abnormal preponderance of sialated apoC-III (5, 6). ApoC-III is synthesized mainly by the liver (7, 8) and participates in the formation of the triglyceride-rich lipoproteins (chylomicrons and VLDL). It has been reported that apoC-III may inhibit lipoprotein lipase, an enzyme responsible for the clearance of triglyceride-rich lipoproteins from the circulation (9–11). In addition, it has been shown that lipoprotein uptake by hepatic tissues is stimulated by apoE and inhibited by apoC-III (12–14). However, the physiological significance of these observations remains to be clarified further. We have previously shown that the human apoC-III gene is located approximately 2.6kb to the 3′ direction of the apoA-I gene and that these two genes are convergently transcribed (15). In addition, we have shown that a C–G transversion in the 3′ noncoding region of apoC-III mRNA generates a polymorphic SacI restriction site (15) which has previously been correlated with hypertriglyceridemia in humans (16).

In this study, we report the isolation and characterization of both SacI-containing (Sac^+ allele) and SacI-lacking (Sac^- allele) cDNA clones from a human liver cDNA library. Analysis of these clones indicated that apoC-III primary translation products contain a 20 amino acid N-terminal extension and that both Sac^+ and Sac^- alleles code for the same apoC-III primary translation product. Finally, we report four amino acid differences between the previously determined (1, 2) and the DNA-derived mature apoC-III sequence.

MATERIALS AND METHODS

Isolation of ds cDNA apoC-III clones

The previously constructed (17) adult human liver cDNA library was plated on 100-mm petri dishes at a

bacterial density of 1000 colonies per dish. After growth and chloramphenicol amplification, the colonies were transferred to nitrocellulose filters as previously described (18). These filters were baked and prehybridized to probes prepared by nick translation (19).

Plasmid DNA preparation and DNA sequencing analysis

Recombinant plasmid was prepared from a 1-liter growth of the appropriate clones by the alkaline lysis method (20) with the modification we described elsewhere (21). Purified recombinant plasmids were digested with PstI to release the cloned DNA inserts. The inserts were isolated by electrophoresis in 1% low melt agarose gel (Bio-Rad) and purified by phenol extraction followed by ethanol precipitation. These DNA inserts were used for DNA sequencing analysis according to either the Maxam and Gilbert Method (22) or the method of Sanger, Nicklen, and Coulson (23) with the aid of M13 cloning-sequencing system (24) using reverse transcriptase instead of the Klenow fragment of *E. coli* DNA polymerase (S. K. Karathanasis, unpublished data).

Genomic blotting methods

Chromosomal DNA was prepared from peripheral lymphocytes as we have described elsewhere (25). Genomic blotting analysis was carried out using the Southern method (26).

RESULTS

Isolation and characterization of apoC-III cDNA clones

We have previously reported the isolation and characterization of a human genomic clone (λapoA-I #6) which contains both the apoA-I and apoC-III genes (21). Further characterization of this clone indicated that it contains the entire apoC-III gene, a restriction map of which is diagrammatically shown in **Fig. 1.** The approximately 2.2kb BamHI-EcoRI fragment (Fig. 1) was subcloned from λ apoA-I #6 in pUC9 (27) and, after purification and ^{32}P-labeling by nick translation, was used to screen the previously constructed human liver cDNA library (17). Screening of approximately 40,000 cDNA clones yielded six positive clones, which upon rescreening with the labeled insert of clone pCIII-606 (15), indicated that all these six clones contain sequences relevant to apoC-III cDNA. Three of these positive clones, pCIII-607, pCIII-655-1, and pCIII-655-2, were grown, recombinant plasmids were isolated, and the plasmid inserts were purified as described in Methods.

Restriction endonuclease mapping of these inserts was carried out and the resulting restriction maps are diagrammatically shown in **Fig. 2.** The pCIII-607 clone and the previously reported (15) pCIII-606 clone both contain inserts with the SacI restriction endonuclease site. However, neither the pCIII-655-1 clone nor the pCIII-655-2 clone contain inserts with this SacI restriction site. This observation implied that the individual from whom this liver cDNA library had been constructed was heterozygous for this SacI restriction site. To verify this, chromosomal DNA was prepared from another aliquot of the same liver used to construct the cDNA library and, after digestion with SacI, electrophoresis, blotting, and hybridization with the apoA-I cDNA probe pAI-101 (28), produced 5.7kb, 4.2kb, and 3.2kb hybridization bands (data not shown). This is compatible with the previously reported (16) SacI heterozygous pattern found to be more common in hypertriglyceridemic patients.

Nucleotide sequences and the derived amino acid sequence of the DNA inserts in apoC-III cDNA clones

The purified inserts of clones pCIII-607 and pCIII-655-2 were subjected to nucleotide sequencing analysis as described in Methods. The nucleotide sequence strategy is outlined in Fig. 2, and the resulting nucleotide sequences, as well as the derived amino acid sequence, are shown in **Fig. 3.** The size of the cDNA insert found in the longest cDNA clone (pCIII-655-2) is 519 b.p. (Fig. 2), which is compatible with the size of approximately 700 nucleotides of the apoC-III mRNA as determined by Northern blotting analysis (data not shown). Inspection of

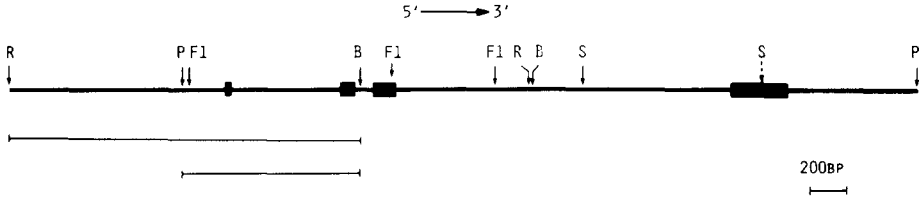

Fig. 1 Restriction map of the human apoC-III gene. Regions of the gene present in the mature apoC-III mRNA are shown by solid boxes. Lines under the restriction map show the segments of apoC-III gene used to prepare radioactively labeled probes (see text). Restriction sites are indicated as follows: R) EcoRI, B) BamHI, P) PstI, S) SacI, F1) HinfI. The polymorphic (see text) SacI restriction site is shown by a dotted-line arrow.

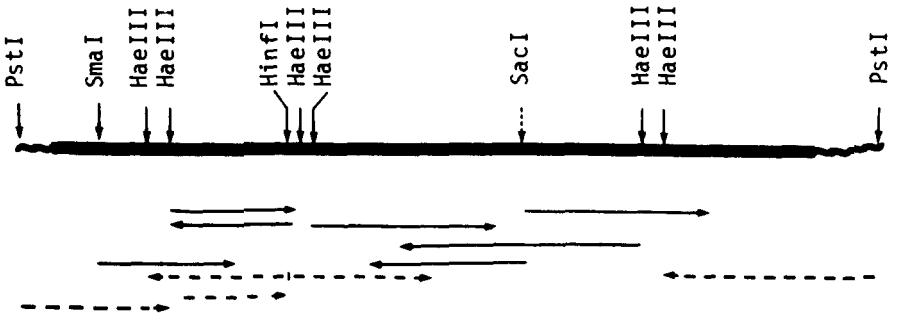

Fig. 2 Restriction map and sequencing strategy of the cDNA inserts in pCIII-607 and pCIII-655-2. Only relevant restriction sites are shown. Horizontal arrows indicate the direction and extent of DNA sequence determinations. Broken or intact-line arrows show nucleotide sequences obtained by the Maxam and Gilbert (22) and Sanger et al. (23) methods, respectively. Wavy lines represent the GC tails flanking the clone inserts.

the DNA sequence to the 5' direction of the apoC-III +1 amino acid revealed the presence of a 20 amino acid N-terminal extension (Fig. 3). This apoC-III N-terminal extension begins with methionine (residue −20), has arginine at residue −17, and ends with alanine at residues −3 and −1. These features are typical of previously reported signal peptide sequences (29–32) and indicate that this 20 amino acid long apoC-III N-terminal extension represents the human apoC-III signal peptide. In agreement with this observation, the recently reported rat apoC-III signal peptide sequence (33) differs only at residue −16 (methionine-rat:valine-human) from the DNA-derived sequence of human apoC-III signal peptide (Fig. 3).

The amino acid sequence of mature apoC-III has been previously reported (1, 2). Comparison between this previously reported amino acid sequence and the DNA-derived apoC-III amino acid sequence indicates differences in four residues. More specifically, at residue positions +32, +33, +37, and +39, the DNA sequence predicts Glu, Ser, Gln, Ala, while the previously reported amino acid sequence specifies Ser, Gln, Ala, Gln, respectively (Fig. 3). Although it is conceivable that these amino acid differences may reflect apoC-III polymorphisms, it is clear from our genomic blotting examination (see below) that at least in the case of six different individuals examined, the apoC-III +32–+33 residues are compatible with the nucleotide-derived amino acid sequence.

Finally, as mentioned above, pCIII-607 and pCIII-655-2 contain cDNA inserts representing the transcriptional products of two forms of the apoC-III gene, one containing and one lacking a SacI site. There are three nucleotide differences between the cDNA of these two alleles consisting of T-C transition, G-C (SacI site polymorphism) transversion, and a G-T transversion at nucleotide positions 132, 370, and 401, respectively (Fig. 3). However, none of these differences has an effect on the apoC-III primary amino acid sequence (Fig. 3).

Genomic blotting examination of apoC-III amino acid residues +32 Glu–+33 Ser in normal and hypertriglyceridemic individuals

As has been mentioned above, there are four residue differences between the previously described and the nucleotide-derived apoC-III amino acid sequences. Inspection of these differences indicated that apoC-III residues +32 Glu, +33 Ser predicted by the DNA sequence are derived by the codons GAG and TCC, respectively. This GAGTCC sequence includes the recognition site for the restriction endonuclease HinfI (GANTC). On the other hand, apoC-III residues +32 Ser, +33 Glu reported previously cannot be derived by codons that include this HinfI site. Fig. 1 shows that this +32 Glu–+33 Ser HinfI site in λapoA-I #6 is flanked by HinfI sites 1.15kb and 0.6kb to the 5' and 3' directions, respectively. In addition, Fig. 1 shows that this +32 Glu–+33 Ser HinfI site is flanked to the 5' direction by a BamHI which occurs at a distance of approximately 0.2kb.

Therefore, HinfI or HinfI + BamHI genomic blotting analysis of human DNA, using as a probe the lkb PstI-BamHI fragment indicated in Fig. 1, will be expected to produce 1.75kb and 0.95kb or 1.15kb and 0.95kb hybridization bands for absence or presence of this +32–+33 HinfI site, respectively. **Fig. 4** shows HinfI and HinfI-BamHI genomic blotting analysis of four normal and two hypertriglyceridemic individuals. As can be seen in all cases, HinfI and HinfI-BamHI digestion produced 1.15kb and 0.95kb hybridization bands, respectively. These results indicate that at least in the case of the individuals examined, the apoC-III residues +32–+33 are consistent with the nucleotide-derived apoC-III amino acid sequence.

DISCUSSION

ApoC-III, a 79 amino acid polypeptide, is a major

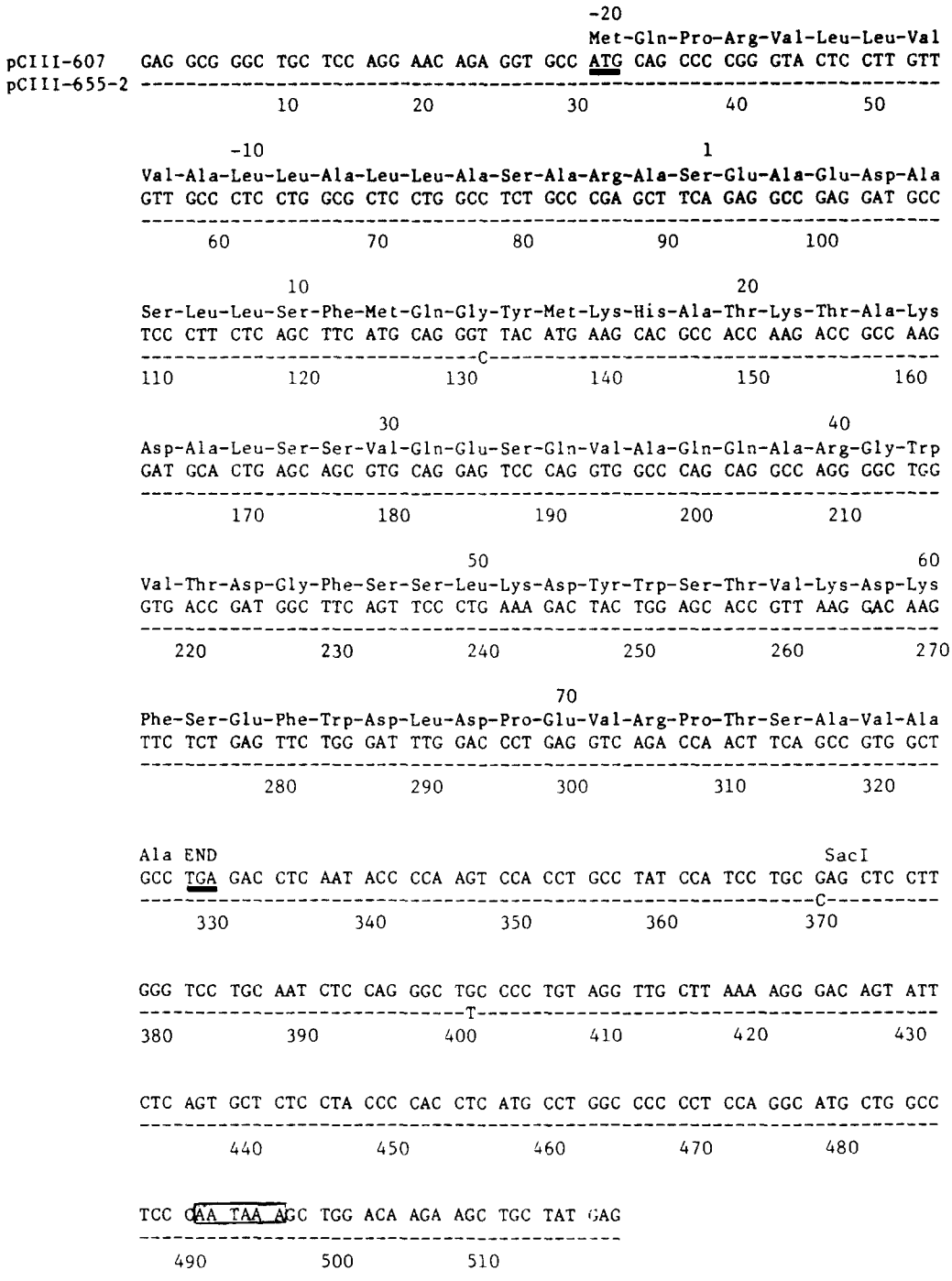

Fig. 3 The nucleotide and the derived amino acid apoC-III sequences of pCIII-607 are shown. Initiation and termination codons are underlined. The polyadenylation signal is boxed. Positive and negative numbers above the animo acid sequence indicate mature and signal peptide sequences, respectively. Nucleotide residues are numbered below the nucleotide sequences. Differences between pCIII-607 and pCIII-655-2 (dotted lines) are indicated. The polymorphic SacI restriction site is indicated.

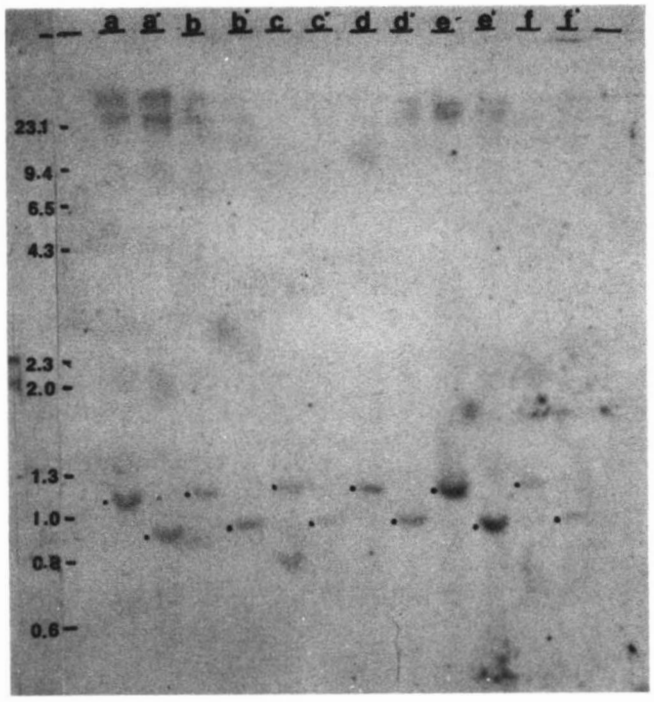

Fig. 4 Genomic blotting analysis of the +32 through +33 apoC-III residues in normal and hypertriglyceridemic individuals. Chromosomal DNA prepared from normal (a, b, c, d) or hypertriglyceridemic individuals (e, f), digested with either HinfI (a, b, c, d, e, f) or HinfI plus BamHI (a', b', c', d', e', f'), electrophoresed on 1% agarose gel, transferred on a nitrocellulose filter, and hybridized with the ^{32}P-labeled PstI-BamHI DNA fragment (see Fig. 1). (Note the approximately 0.8-kb band in lane c is due to background hybridization.) The resulting autoradiogram is shown. Molecular weight standards are HindIII-digested λDNA and HaeIII-digested ϕX174 DNA (New England Biolabs).

protein component of very low density lipoprotein (VLDL) and some reports have indicated that it may be involved in inhibition of lipoprotein lipase (LPL), an enzyme responsible for the clearance of VLDL and chylomicrons from the circulation (8–10). It has been previously (16) reported that a SacI restriction site polymorphism in the apoA-I–apoC-III gene locus may be associated with the development of hypertriglyceridemia in humans. In addition, we have recently shown that this SacI site polymorphism is due to a G-C transversion in the 3' noncoding region of apoC-III mRNA (15). In this report, we show that the apoC-III-amino acid sequence derived from the nucleotide sequence of the SacI site-containing apoC-III cDNA is identical to the amino acid sequence derived from the SacI site-lacking apoC-III cDNA. The nucleotide sequence of apoC-III cDNA obtained from the Sac$^-$ clone has a T-C transition at nucleotide 132 and one G-C and one G-T transversion at nucleotides 370 and 401, respectively, as compared to the Sac$^+$ clone. Compared to the DNA sequence of another recently reported Sac$^-$ apoC-III clone, our Sac$^-$ apoC-III clone has a T-C transition and a G-T transversion at nucleotides 132 and 401, respectively (34). However, the amino acid sequence of all these apparently different apoC-III alleles is identical (34). These results indicate that the presence of this SacI site polymorphism in the genomic DNA of hypertriglyceridemic patients may not be directly related to alterations in apoC-III structure. Therefore, association of this SacI polymorphic site with hypertriglyceridemia may indicate alterations in apoC-III gene expression or regulation of expression. Alternatively, this polymorphism may be associated with alterations in the closely linked apoA-I gene or some other yet unidentified nearby gene.

The amino acid sequence derived from the nucleotide sequence of apoC-III (Fig. 3) shows that the protein contains a 20 amino acid N-terminal extension conforming with the general properties of signal peptides (29–32). This result, therefore, indicates that apoC-III is synthesized as a preprotein, which upon cleavage of the signal peptide, is converted to the mature form occurring in human plasma. In addition, the nucleotide-derived mature apoC-III primary structure differs in four residues (+32, +33, +37, +39) from the previously reported apoC-III amino acid sequence (Fig. 3). It is interesting that the previously reported sequence was derived from apoC-III purified from the plasma of a hypertriglyceridemic patient. It is, therefore, conceivable that the previously determined sequence may represent the product of an apoC-III gene allele responsible for, or linked to, the development of hypertriglyceridemia. To examine this possibility, we devised a method by which the presence of this allele can be assayed by genomic blotting analysis of chromosomal DNA isolated from peripheral lymphocytes of various individuals (see Results). Using this method, we examined four normal and two hypertriglyceridemic individuals, and the results indicated that none of these individuals has the apoC-III gene allele reported previously (Fig. 4). Clearly, this methodology should facilitate population and/or family studies to examine the possible linkage of this previously reported apoC-III allele with the development of hypertriglyceridemia.

This work was supported by grants from the National Institutes of Health (HL32032, HL32354, and HL33952), the March of Dimes Birth Defects Foundation (1-817), the American Heart Association (83-963), and the Massachusetts Affiliate of the American Heart Association (13-517-845). Dr. Jan L. Breslow and Dr. Vassilis I. Zannis are Established Investigators of the American Heart Association. We would like to thank Ms. Hilary Williams, Ms. Carolyn Frazer, Ms. Elizabeth Salmon, and Ms. Lorraine Duda for their expert assistance.

Manuscript received 12 July 1984.

REFERENCES

1. Brewer, H. B., Jr., R. Shulman, P. Herbert. R. Ronan, and K. Wehrly. 1974. The complete amino acid sequence of alanine apolipoprotein (apoC-III), an apolipoprotein from

human plasma very low density lipoproteins. *J. Biol. Chem.* **249:** 4975–4984.
2. Shulman, R. S., P. N. Herbert, D. S. Fredrickson, K. Wehrly, and H. B. Brewer, Jr. 1974. Isolation and alignment of the tryptic peptides of alanine apolipoprotein, an apolipoprotein from human plasma very low density lipoprotein. *J. Biol. Chem.* **249:** 4969–4974.
3. Brown, W. V., R. I. Levy, and D. S. Fredrickson. 1970. Further characterization of apolipoproteins from the human plasma very low density lipoproteins. *J. Biol. Chem.* **245:** 6588–6594.
4. Zannis, V. I., V. Rooney, P. Fraser, and J. L. Breslow. 1983. Use of electrophoretic techniques for detection of molecular defects of lipoprotein metabolism. *In* CRC Handbook of Electrophoresis. L. A. Lewis, editor. **3:** 319–348.
5. Kashyap, M. L., B. A. Hynd, K. Robinson, and P. S. Gartside. 1981. Abnormal preponderance of sialylated apolipoprotein C-III in triglyceride-rich lipoproteins in type V hyperlipoproteinemia. *Metabolism.* **30:** 111–118.
6. Kashyap, M. L., L. S. Srivastava, B. A. Hynd, P. S. Gartside, and G. Perisutti. 1981. Quantitation of human apolipoprotein C-III and its subspecies by radioimmunoassay and analytical isoelectric focusing: abnormal plasma triglyceride-rich lipoprotein apolipoprotein C-III subspecie concentrations in hypertriglyceridemia. *J. Lipid Res.* **22:** 800–810.
7. Zannis, V. I., D. Kurnit, and J. L. Breslow. 1982. Hepatic apo-A-I and apo-E and intestinal apo-A-I are synthesized in precursor isoprotein forms by organ cultures of human fetal tissues. *J. Biol. Chem.* **257:** 536–544.
8. Wu, A. L., and H. G. Windmueller. 1979. Relative contribution by liver and intestine to individual plasma apolipoproteins in the rat. *J. Biol. Chem.* **254:** 7316–7322.
9. Brown, W. V., and M. L. Baginsky. 1972. Inhibition of lipoprotein lipase by an apoprotein of human very low density lipoprotein. *Biochem. Biophys. Res. Commun.* **46:** 375–382.
10. Krauss, R. M., P. M. Herbert, R. I. Levy, and D. S. Fredrickson. 1973. Further observations on the activation and inhibition of lipoprotein lipase by apolipoproteins. *Circ. Res.* **33:** 403–411.
11. Chung, J., and A. M. Scanu. 1977. Isolation, molecular properties and kinetic characterization of lipoprotein lipase from rat heart. *J. Biol. Chem.* **252:** 4204–4209.
12. Windler, E., Y. Chao, and R. J. Havel. 1980. Determinants of hepatic uptake of triglyceride-rich lipoproteins and their remnants in the rat. *J. Biol. Chem.* **255:** 5475–5480.
13. Shelburne, F., J. Hanks, W. Meyers, and S. Quarfordt. 1980. Effect of apoproteins on hepatic uptake of triglyceride emulsions in the rat. *J. Clin. Invest.* **65:** 652–658.
14. Quarfordt, S. H., G. Michalopoulos, and B. Schirmer. 1982. The effect of human C apolipoproteins on the in vitro hepatic metabolism of triglyceride emulsions in the rat. *J. Biol. Chem.* **257:** 14642–14647.
15. Karathanasis, S. K., J. McPherson, V. I. Zannis, and J. L. Breslow. 1983. Linkage of human apolipoprotein A-I and C-III genes. *Nature.* **304:** 371–373.
16. Rees, A., C. C. Shoulders, J. Stocks, D. J. Galton, and F. E. Baralle. 1983. DNA polymorphism adjacent to human apoprotein A-I gene: relation to hypertriglyceridemia. *Lancet.* **1:** 444–446.
17. Woods, D. E., A. F. Markham, A. T. Ricker, G. Goldberger, and H. R. Colten. 1982. Isolation of cDNA clones for the human complement protein factor B, a class III major histocompatibility complex gene product. *Proc. Natl. Acad. Sci. USA.* **79:** 5661–5665.
18. Grunstein, M., and D. S. Hogness. 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. USA.* **72:** 3961–3965.
19. Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* **113:** 237–251.
20. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7:** 1513–1523.
21. Karathanasis, S. K., V. I. Zannis, and J. L. Breslow. 1983. Isolation and characterization of the human apolipoprotein A-I gene. *Proc. Natl. Acad. Sci. USA.* **80:** 6147–6151.
22. Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA.* **74:** 560–564.
23. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74:** 5463–5467.
24. Messing, J., R. Crea, and P. H. Seeburg. 1981. A system for shotgun DNA sequencing. *Nucleic Acids Res.* **9:** 309–321.
25. Karathanasis, S. K., R. A. Norum, V. I. Zannis, and J. L. Breslow. 1983. A mutation in the human apo A-I gene locus related to the development of atherosclerosis. *Nature.* **301:** 718–720.
26. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98:** 503–517.
27. Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene.* **19:** 259–268.
28. Karathanasis, S. K., V. I. Zannis, and J. L. Breslow. 1983. A DNA insertion in the apolipoprotein A-I gene is related to apolipoprotein A-I, apolipoprotein C-III and high density lipoprotein deficiencies in patients with premature atherosclerosis. *Nature.* **305:** 823–825.
29. Lingappa, V. R. 1979. *In* Symposium of the Society of Experimental Biology (Great Britain). C. R. Hopkins and C. J. Duncan, editors. Cambridge University Press, London. **33:** 9–36.
30. Habener, J. F., M. Rosenblatt, B. Kemper, H. M. Kronenberg, A. Rich, and J. T. Potts, Jr. 1978. Pre-proparathyroid hormone: amino acid sequence, chemical synthesis and some biological studies of the precursor region. *Proc. Natl. Acad. Sci. USA.* **75:** 2616–2620.
31. Inouye, M., and S. Halegoua. 1980. Secretion and membrane localization of proteins in *Escherichia coli*. *CRC Crit. Rev. Biochem.* **7:** 339–371.
32. Austen, B. M. 1979. Predicted secondary structures of amino-terminal extension sequences of secreted proteins. *FEBS Lett.* **103:** 308–313.
33. Blaufuss, M. C., J. I. Gordon, G. Schonfeld, A. W. Strauss, and D. H. Alpers. 1984. Biosynthesis of apolipoprotein C-III in rat liver and small intestinal mucosa. *J. Biol. Chem.* **259:** 2452–2456.
34. Sharpe, C. R., A. Sidoli, C. S. Shelley, M. A. Lucero, C. C. Shoulders, and F. E. Baralle. 1984. Human apolipoproteins AI, AII, CII, and CIII. cDNA sequences and mRNA abundance. *Nucleic Acids Res.* **12:** 3917–3932.