Amino acid sequence of human plasma apolipoprotein C-III from normalipidemic subjects

Ashok V. Hospattankar, H. Bryan Brewer, jr, Rosemary Ronan and Thomas Fairwell*

Molecular Disease Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received 23 December 1985

The complete amino acid sequence of human plasma apolipoprotein C-III (apoC-III) isolated from normal subjects is described. ApoC-III is a linear polypeptide chain of 79 amino acids. Tryptic digestion of intact apoC-III produced 5 major peptides, while tryptic digestion of the citraconylated protein yielded two peptides. The complete amino acid sequence of apoC-III was determined by the automated Edman degradation of the intact protein as well as the various tryptic peptides. Phenylthiohydantoin amino acids were identified by high-performance liquid chromatography and chemical ionization mass spectrometry. The amino acid sequence of apoC-III isolated from normalipidemic subjects is identical to the apoC-III sequence derived from the cDNA sequence and differs at 4 positions from the previously reported sequence of apoC-III derived from a patient with type V hyperlipoproteinemia.

Amino acid sequence *Apolipoprotein C-III*

1. INTRODUCTION

The C apolipoproteins, apoC-I, apoC-II and apoC-III, are present primarily in the HDL of normal subjects and in the chylomicron-VLDL of hyperlipoproteinemic patients. These low-M_r proteins have been proposed to have specific physiological roles in lipoprotein metabolism. ApoC-I binds to phospholipids and has been reported to increase the activity of lecithin-cholesterol acyltransferase in vitro [1,2]. ApoC-II has been shown to activate lipoprotein lipase in vitro [3,4] and a deficiency of apoC-II is associated with severe hypertriglyceridemia [5]. ApoC-III has been proposed to have an important role in the metabolism of triglyceride rich lipoproteins, since it inhibits the activation of lipoprotein lipase by apoC-II and prevents the uptake of lipoproteins containing apoE by liver cells [6–10]. Human apoC-III is a glycoprotein and occurs in three isoforms, containing 0, 1 or 2 molecules of sialic acid/molecule of protein [11,12]. The primary amino acid sequence of apoC-III isolated from a patient with type V hyperlipoproteinemia has been reported previously and is a linear peptide of 79 amino acids, with a carbohydrate moiety attached to the threonine at position 74 [13]. Qualitative studies based on electrophoretic mobility and immunological properties suggested that there is no difference between the apoC-III from normal individuals and those of patients with familial hyperlipoproteinemia [12]. However, the recently reported [14] protein sequence derived from the cDNA sequence for apoC-III differs from the amino acid sequence at 4 positions. The lipid levels of the subjects in this study were not reported and is is not clear whether the liver used came from normal subjects or a patient with hyperlipoproteinemia. We have isolated apoC-III from normalipidemic individuals and in

* To whom correspondence should be addressed

Abbreviations: apo, apolipoprotein; VLDL, very low density lipoproteins; HDL, high density lipoproteins; HPLC, high-performance liquid chromatography; PTH, phenylthiodantoin

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

this report the complete amino acid sequence of normal apoC-III is described. The sequence was determined by automated Edman degradation of the intact protein, and the tryptic peptides of intact and citraconylated apoC-III.

2. MATERIALS AND METHODS

The apoC-III used in these studies was isolated from the plasma of normalipidemic volunteers. The isolation and purification procedures have been described in detail [12]. In brief, HDL or VLDL were isolated by preparative ultracentrifugation and delipidated with chloroform-methanol (2:1). The C apolipoproteins were initially purified on a Sephadex G-200 column in 0.2 M Tris-HCl in 6 M urea (pH 8.0). The fraction containing the C apolipoproteins, as well as the different isoforms of apoC-III, was separated on a DE-52 (Whatman) column utilizing a linear gradient of Tris-HCl in 6 M urea (0.01–0.165 M, pH 8.0) [12]. ApoC-III$_1$, the isoform containing 1 mol sialic acid/mol protein, was used here.

2.1. *Tryptic digestion of apoC-III*

Tryptic peptides were generated by digestion of apoC-III with trypsin (Worthington, 100:1, w/w) for 2.5 h at 37°C in 0.2 M ammonium bicarbonate (pH 8.5). In a separate study the cleavage of apoC-III by trypsin was limited to the arginines by modification of apoC-III with citraconic anhydride at a 100-fold excess over the amino groups at pH 9.6 followed by cleavage of the lysine-modified apoC-III with trypsin (50:1, w/w) for 2.5 h at 37°C. The resultant peptide fragments were deblocked by addition of acetic acid (pH 3.0).

2.2. *Isolation of peptides*

Tryptic peptides of intact and citraconylated apoC-III were separated on reversed-phase HPLC using a Dupont 8800 HPLC system and a Waters C-18 μBondapack column. The peptides were separated by a programmed binary-solvent gradient system using 0.1% aqueous trifluoroacetic acid and acetonitrile as the solvents. The isolated peptides wre monitored for purity by amino acid analysis (24 h at 108°C in constant boiling HCl) using a Beckman 6300 analyzer, thin layer chromatography as well as amino-terminal sequence analysis as described in earlier reports [12,13].

2.3. *Edman degradation*

Intact apoC-III and the isolated tryptic peptides were sequenced by automated Edman procedure [15] using a modified Beckman 890B sequencer, with a 0.1 M quadrol program [16]. Polybrene was used as a carrier in all degradations [17]. PTH derivatives were quantitatively identified by HPLC [18] on a Dupont-Zorbax ODS column using a Dupont 8800 LC system equipped with a Perkin-Elmer Chrom-2 data system and an ISS-100 autosampler. All PTH derivatives were also identified by chemical ionization mass spectrometry using a Finnigan 4510 mass spectrometer and isobutane as the reagent gas [19].

3. RESULTS

The amino acid compositions of intact normal apoC-III, tryptic as well as citraconylated tryptic peptides, are shown in table 1. The amino acid composition of the tryptic peptides is identical to that of the apoC-III from a hyperlipoproteinemia subject previously reported [12].

The complete amino acid sequence of apoC-III determined by automated Edman degradation of the intact protein and the various tryptic peptides is illustrated in fig.1. The sequence of the initial 43 residues from the amino-terminal end of the molecule was obtained by automated Edman degradation of intact apoC-III (fig.1, large arrows). Serine was identified as the amino-terminal residue and the first arginine residue was located at step 40 as reported in [13]. Since the amino acid composition revealed the presence of 6 lysines and 2 arginines, as many as 9 peptides would be expected from tryptic digestion. However, no cleavage occurred at the arginyl-proline bond (residues 72–73) and only partial cleavages occurred at 3 lysine residues (residues 21, 24 and 60) resulting in only 5 major peptides and 2 citraconylated tryptic peptides. The presence of sialic acid has been reported to influence the extent of tryptic cleavage in apoC-III [12,13]. Since the small quantities of the 3 short peptides would have provided only redundant amino acid sequence information, no attempt was made to isolate these peptides. The 5 major peptides generated from the

Table 1

Amino acid composition of native and peptide fragments of normal apoC-III

Amino acids	ApoC-III	TP-1	TP-2	TP-3	TP-4	TP-5	CT-1	CT-2
Aspartic acid	6.95(7)	1.13(1)	1.46(1)	1.17(1)	1.05(1)	2.59(3)	2.21(2)	4.98(5)
Threonine	5.01(5)		2.12(2)	1.03(1)	0.90(1)	1.02(1)	2.06(2)	2.94(3)
Serine	10.43(11)	2.86(3)	2.94(3)	2.01(2)	1.24(1)	2.13(2)	5.38(6)	5.07(5)
Glutamic acid	9.73(10)	2.75(3)	4.38(5)			2.15(2)	7.29(8)	2.41(2)
Proline	2.18(2)					1.91(2)		1.86(2)
Glycine	3.10(3)	1.14(1)		1.94(2)			1.21(1)	2.34(2)
Alanine	10.02(10)	1.92(2)	4.67(5)			2.96(3)	7.24(7)	3.26(3)
Valine	5.54(6)		1.93(2)	0.94(1)	1.03(1)	1.84(2)	2.01(2)	3.58(4)
Methionine	1.83(2)	1.46(2)					1.49(2)	
Leucine	5.20(5)	2.01(2)	1.03(1)	1.18(1)		1.21(1)	3.31(3)	2.31(2)
Tyrosine	2.14(2)	1.03(1)			0.87(1)		1.15(1)	1.01(1)
Phenylalanine	4.09(4)	1.08(1)		1.12(1)		2.23(2)	1.23(1)	3.01(3)
Histidine	1.00(1)		0.91(1)				0.98(1)	
Lysine	6.15(6)	1.07(1)	2.31(2)	1.01(1)	0.89(1)	0.79(1)	3.05(3)	2.98(3)
Arginine	1.92(2)		1.01(1)			0.82(1)	0.92(1)	0.80(1)
Tryptophan	3.30(3)[a]			ND(1)	ND(1)	ND(1)		ND(1)

[a] Determined spectrophotometrically

All values are expressed as molar ratios of the constituent amino acids following a 24 h hydrolysis at 108°C. Integral values in parentheses are numbers of residues/mol based on sequence data

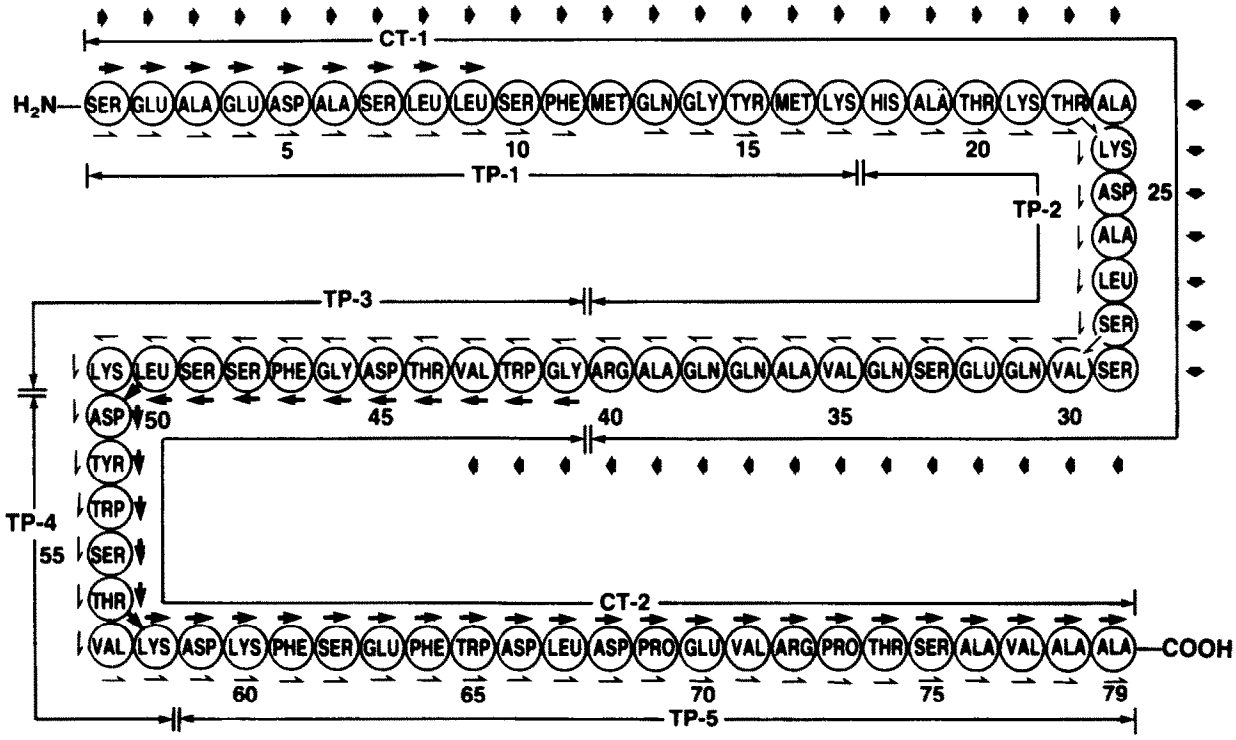

Fig. 1. The complete amino acid sequence of human plasma apolipoprotein C-III. The large arrows denote automated Edman degradations performed on intact apoC-III. The half arrows refer to Edman degradations of isolated tryptic peptides and the long arrows indicate the degradations on the tryptic peptides after citraconylation of apoC-III.

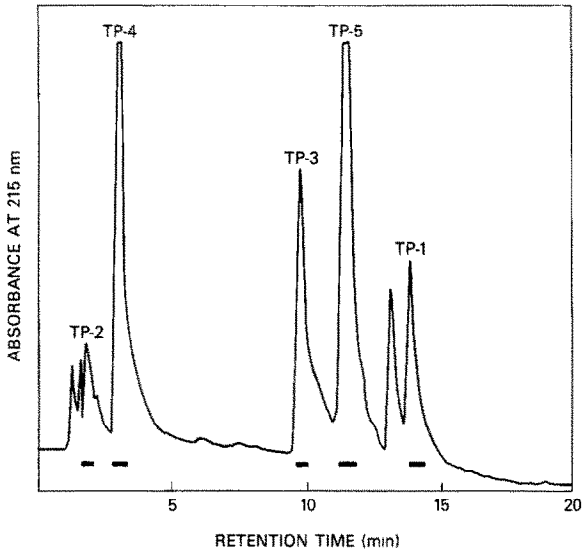

Fig.2. The elution profile of tryptic peptides on HPLC. Bar indicates peptide fraction pooled for sequence analysis. A programmable binary-solvent gradient system (25–60% acetonitrile) with aqueous trifluoroacetic acid and acetonitrile as solvents was used for the separation of peptides. The peak at 13.6 min was shown to be a mixture of TP-1 and TP-5 by sequence analysis.

tryptic digestion of intact apoC-III were separated by HPLC (fig.2). These peptides are designated TP-1 (17 amino acids), TP-2 (23 amino acids), TP-3 (11 amino acids), TP-4 (7 amino acids) and TP-5 (21 amino acids) (table 1). The complete sequences of each of these peptides were obtained as follows:

Peptide TP-1: All 17 residues of TP-1 (50 nmol) were determined by automated Edman degradation (fig.1). Amino acids 1–17 corresponded to residues 1–17 of intact apoC-III. Thus, the sequence data established that TP-1 is from the amino-terminus of apoC-III.

Peptide TP-2: The complete sequence of the 23 residue peptide (90 nmol) revealed that this peptide is an overlap tryptic fragment of apoC-III with intact lysines at residues 4 and 7 (fig.1). The relatively high yield of this peptide recovered after HPLC purification suggested that very little cleavage occurred at these two lysine residues and more rigorous conditions would be required for complete cleavage at lysine residues 4 and 7. The amino acid sequence of this peptide corresponded to residues 18–40 of the intact apoC-III (fig.1).

Peptide TP-3: The sequence of this 11 residue peptide (35 nmol) was determined by Edman degradation. The first 3 amino acids corresponded to residues 41–43 of the intact molecule (fig.1). This result indicated that tryptic peptide TP-3 followed TP-2 in the sequence of apoC-III and corresponded to residues 41–51.

Peptide TP-4: Automated Edman degradation of peptide TP-4 (25 nmol) established the sequence of this 7 residue peptide. Aspartic acid and lysine were identified as the amino-terminal and carboxy-terminal amino acids, respectively (fig.1).

Peptide TP-5: The total sequence of the 21 amino acid peptide (60 nmol) was obtained by Edman degradation (fig.1). Lysine was identified at step 2 and arginine at step 14, indicating that this peptide was an overlap tryptic peptide. Since the arginine residue is followed by a proline, no cleavage would be expected at the Arg-Pro bond, while the partial cleavage at the Lys-Phe bond is attributed to the presence of the sialic acid containing carbohydrate moiety at step 16 [13]. The threonine at step 16 was not identified by HPLC, however, it was positively identified as dehydrothreonine by mass spectral analysis. Yields were reduced at the end of the degradation due to the several hydrophobic residues located at the carboxy-terminal region of the peptide and the loss of peptide from the sequencer cup (fig.1).

The combined results obtained from the Edman degradation of the intact apoC-III and tryptic peptides TP-1, TP-2, and TP-3 established the complete amino acid sequence of residues 1–51.

Peptides CT-1 and CT-2: To ascertain the alignment of peptides TP-3, TP-4, and TP-5, apoC-III was cleaved by trypsin at the arginine residue. Specific cleavage at the arginine residues was achieved by modification of the lysines by citraconylation and tryptic digestion of the modified protein. Even though there are two arginines in the apoC-III molecule, one arginine (residue 14 in peptide TP-5) is not amenable to tryptic digestion since it is followed by a proline in the sequence. As a result, only two peptides, CT-1

(40 amino acids) and CT-2 (39 amino acids) were obtained by tryptic digestion of citraconylated apoC-III and were separated by HPLC.

The amino acid sequence of the first 9 residues of CT-1 was the same as that of intact apoC-III and indicated that this peptide was the amino-terminal peptide (fig.1). The complete sequence of peptide CT-2 was determined by automated Edman degradation (fig.1). Glycine was identified as the amino-terminal residue. The first 11 residues corresponded to the total sequence of peptide TP-3, and residues 12–18 corresponded to the sequence determined for TP-4. These results confirmed the sequence of each peptide and established the alignment of peptides TP-3, TP-4, and TP-5 in apoC-III (fig.1). The threonine at position 34 in the degradation of peptide CT-2 was identified only by mass spectrometry as the dehydrothreonine derivative as discussed above.

The combined results from the Edman degradations of intact apoC-III, the tryptic, and citraconylated tryptic peptides established the complete sequence of normal apoC-III as shown in fig.1.

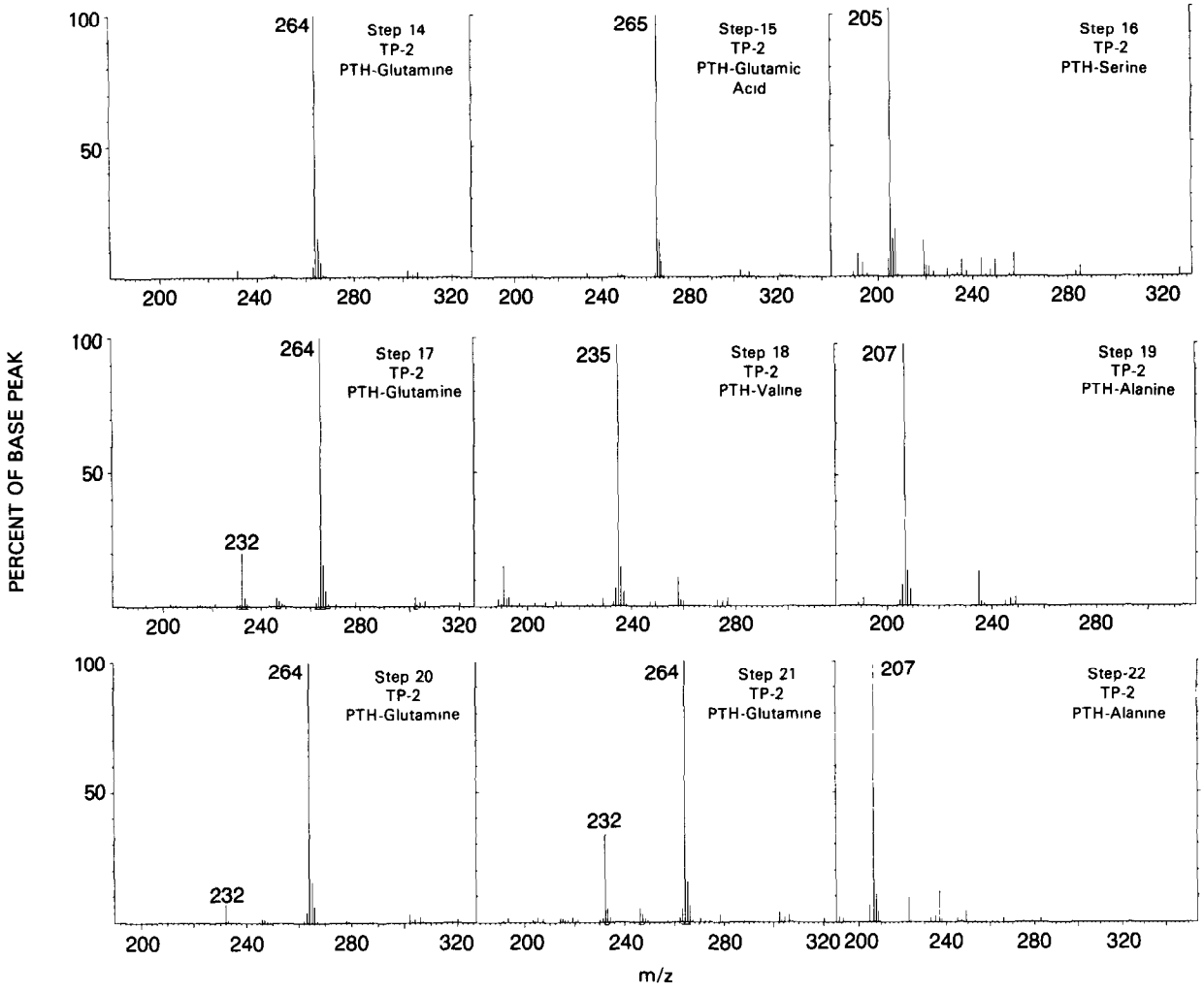

Fig.3. Identification of phenylthiohydantoin derivatives of apoC-III by chemical ionization mass spectrometry. Chemical ionization mass spectra of PTH derivatives (residues 31–39) obtained during the automated Edman degradation of tryptic peptide TP-2 (residues 14–22) are illustrated.

4. DISCUSSION

ApoC-III has been shown previously to be a glycoprotein containing 1 mol each of galactosamine and galactose/mol protein [12]. Three polymorphic forms of this protein have been identified containing 0, 1 or 2 mol sialic acid/mol protein [12,13]. The carbohydrate moiety is attached to the protein through *O*-glycosidic linkage at the threonine at position 74 [13].

Human apoC-III isolated from HDL of normal individuals is composed of a linear polypeptide chain of 79 amino acids with serine as the amino-terminal and alanine as the carboxy-terminal residue. The amino acid compositions of normal apoC-III obtained by acid hydrolysis of the intact molecule and the various peptide fragments are in complete agreement with the data obtained from the amino acid sequence determined by automated Edman degradation of intact apoC-III and tryptic as well as citraconylated tryptic peptide (fig.1, table 1). The threonine residue at position 74, which contains the carbohydrate residue was identified by mass spectral analysis in peptides TP-5 and CT-2.

The amino acid sequence of apoC-III presented here differs at 4 positions from the previously published sequence for apoC-III, which was isolated from a patient with type V hyperlipoproteinemia. These differences occur at positions 32, 33, 37 and 39 in the sequence. In the previous analysis, the amino acids reported at these positions were Ser, Gln, Ala and Gln, respectively. Here, residue 32 was identified as Glu, residue 33 as Ser, residue 37 as Gln and residue 39 as Ala. The chemical ionization mass spectra of all amino acids from residue 31 to 39 (taken from the sequence analysis of the tryptic peptide TP-2, residues 14–22) are illustrated in fig.3. These sequence data are in complete agreement with the recently reported structure of apoC-III derived from the cDNA sequence of human apoC-III [14].

It is of interest that structural variants have been reported previously for several other human plasma apolipoproteins. Variation in primary structure has been reported for apoA-I, and designated apoA-I_{Milano} [20–22], apoA-$I_{Giessen}$ [23], apoA-$I_{Marburg}$ [23] and apoA-$I_{Munster}$ [24]. Patients with apoA-I_{Milano} and apoA-$I_{Marburg}$ have mild hypertriglyceridemia and reduced levels of HDL [20,23]. Polymorphism due to variation in the primary structure of apoE has also been reported [25–27]. In addition, multiple substitution of several amino acids has been identified in a variant of apoE [28,29].

The structural variation in apoC-III in the previous study [13] may be of pathophysiological importance in the development of hyperlipoproteinemia and further studies will be required to identify the occurrence of structural mutants of apoC-III in other patients with hyperlipoproteinemia. The elucidation of the primary structure of apoC-III derived from normalipidemic subjects now enables the chemical synthesis of intact apoC-III and its various fragments which may be used to analyze the role of apoC-III in modulating lipoprotein lipase activity and the effect of structural mutants of apoC-III on the hepatic uptake of lipoproteins and the development of dyslipoproteinemias.

REFERENCES

[1] Soutar, A.K., Garner, C.W., Baker, H.N., Sparrow, J.T., Jackson, R.L., Gotto, A.M. and Smith, L.C. (1975) Biochemistry 14, 3057–3064.
[2] Albers, J.J., Lin, J. and Roberts, C.P. (1979) Artery 5, 61–75.
[3] LaRosa, J.C., Levy, R.I., Herbert, P., Lux, S.E. and Fredrickson, D.S. (1970) Biochem. Biophys. Res. Commun. 41, 57–62.
[4] Havel, R.J., Shore, J.G., Shore, B. and Bier, D.M. (1970) Circ. Res. 27, 597–600.
[5] Breckenridge, W.C., Little, J.A., Steiner, G., Chow, A. and Poapst, M. (1978) N. Engl. J. Med. 298, 1265–1272.
[6] Miller, N.E., Rao, S.N., Alaupovic, P., Noble, N., Slack, J., Brunzell, J.D. and Lewis, B. (1981) Eur. J. Clin. Invest. 11, 69–76.
[7] Brown, W.V. and Baginsky, M.L. (1972) Biochem. Biophys. Res. Commun. 46, 375–381.
[8] Krauss, R.M., Herbert, P.N., Levy, R.I. and Fredrickson, D.S. (1973) Circ. Res. 33, 403–411.
[9] Windler, E., Chao, Y. and Havel, R.J. (1980) J. Biol. Chem. 255, 5473–5480.
[10] Kashyap, M.L., Srivastava, L.S., Hynd, B.A., Cartside, P. and Perisutti, G. (1981) J. Lipid Res. 22, 800–810.
[11] Albers, J.J. and Scanu, A.M. (1971) Biochim. Biophys. Acta 236, 29.
[12] Shulman, R.S., Herbert, P.N., Fredrickson, D.S., Weherly, K. and Brewer, H.B. jr (1974) J. Biol. Chem. 249, 4969–4974.

[13] Brewer, H.B. jr, Shulman, R., Herbert, P., Ronan, R. and Weherly, K. (1974) J. Biol. Chem. 249, 4975-4984.
[14] Sharpe, C.R., Sidoli, A., Shelley, C.S., Lucero, M.A., Shoulders, C.C. and Baralle, F.E. (1984) Nucleic Acids Res. 12, 3917-3932.
[15] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80.
[16] Fairwell, T. and Brewer, H.B. jr (1979) Anal. Biochem. 99, 242-248.
[17] Hunkapiller, M.W. and Hood, L.E. (1978) Biochemistry 17, 2124-2133.
[18] Zimmerman, C., Pisano, J.J. and Appella, E. (1977) Anal. Biochem. 77, 569-573.
[19] Fairwell, T. (1983) Methods Enzymol. 91, 502-511.
[20] Franceschini, G., Sirtori, C.R., Capurso, A., Weisgraber, K.N. and Mahley, R.W. (1980) J. Clin. Invest. 66, 892-900.
[21] Weisgraber, K.N., Bersot, T.P., Mahley, R.W., Franceschini, G. and Sirtori, C.R. (1980) J. Clin. Invest. 66, 901-907.
[22] Weisgraber, K.N., Rall, S.C., Bersot, T.P., Mahley, R.W., Franceschini, G. and Sirtori, C.R. (1983) J. Biol. Chem. 258, 2508-2513.
[23] Utermann, G., Feussner, G., Franceschini, G., Maas, J. and Steinmetz, A. (1982) J. Biol. Chem. 257, 501-507.
[24] Menzel, M.J., Assmann, G., Rall, S.C., Weisgraber, K.N. and Mahley, R.W. (1984) J. Biol. Chem. 259, 3070-3076.
[25] Weisgraber, K.N., Rall, S.C. and Mahley, R.W. (1981) J. Biol. Chem. 256, 9077-9083.
[26] Rall, S.C., Weisgraber, K.N. and Mahley, R.W. (1982) J. Biol. Chem. 257, 4171-4178.
[27] Rall, S.C., Weisgraber, K.N., Innerarity, T.L. and Mahley, R.W. (1982) Proc. Natl. Acad. Sci. USA 79, 4696-4700.
[28] Rees, A., Shoulders, C.C., Stocks, J., Galton, D.J. and Baralle, F.E. (1983) Lancet II, 444-446.
[29] McLean, J.W., Elshourbagy, N.A., Chang, D.J., Mahley, R.W. and Taylor, J.M. (1984) J. Biol. Chem. 259, 6498-6504.