Human apolipoprotein A-II: complete nucleic acid sequence of preproapoA-II

Karl J. Lackner, Simon W. Law and H. Bryan Brewer, jr

Molecular Disease Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20205, USA

Received 25 July 1984

The complete cDNA nucleic acid sequence of preproapolipoprotein (apo) A-II, a major protein constituent of high density lipoproteins, has been determined on clones from a human liver ds-cDNA library. Clones containing ds-cDNA for apoA-II were identified in the human liver ds-cDNA library using synthetic oligonucleotides as probes. Of 3200 clones screened, 4 reacted with the oligonucleotide probes. The DNA sequence coding for amino acids −17 to +17 of apoA-II were determined by Maxam-Gilbert sequence analysis of restriction fragments isolated from one of these clones, pMDB2049. The remainder of the cDNA sequence was established by sequence analysis of a primer extension product synthesized utilizing a restriction fragment near the 5'-end of clone pMDB2049 as primer with total liver mRNA. The apoA-II mRNA encodes for a 100 amino acid protein, preproapoA-II that has an 18 amino acid prepeptide and a 5 amino acid propeptide terminating with a basic dipeptide (Arg-Arg) at the cleavage site to mature apoA-II.

High density lipoprotein *Apolipoprotein A-II* *Oligonucleotide probe* *Lipoprotein metabolism* *Nucleic acid sequence analysis*

1. INTRODUCTION

Human apolipoprotein A-II is a major protein constituent of HDL. ApoA-II, and the other major protein of HDL, apoA-I, constitute approx. 90% of the protein moiety of human HDL [1,2]. HDL has been of particular interest since HDL has been inversely associated with the development of premature cardiovascular disease [3,4]. Both apoA-II as well as apoA-I have been evaluated as potential predictors of cardiovascular disease [5,6].

In man, plasma apoA-II is a 154 amino acid protein composed of two identical chains of 77 amino acid residues linked by a single disulfide bridge at position 6 in the sequence [7]. In nonhuman primates [8], and other species [9,10] apoA-II is a monomer.

Abbreviations: apo, apolipoprotein; HDL, high density lipoproteins; SSC, 1 × SSC is 150 mM NaCl, 15 mM Na citrate

The physiological role of apoA-II in lipoprotein metabolism is not as yet clearly defined. Our laboratory has previously reported that apoA-II activates hepatic lipase in vitro, and may therefore play a role in lipolysis of HDL constituents [11].

Human apoA-II is polymorphic in plasma [12]. Recently apoA-II has been shown by cell free translation to be synthesized as a precursor apolipoprotein, preproapoA-II [13,14]. The partial amino acid sequence of the prepropeptide has been reported by two laboratories. In one report the propeptide was 6 amino acids in length with the amino acid sequence at the cleavage site similar to proapoA-I [13], while in the other study proapoA-II contains a pentapeptide with cleavage of the prosegment after paired basic residues [14]. The hepatoma cell line (Hep G2) has been found to secrete proapoA-II indicating that conversion to the mature protein may occur in plasma [14].

To understand better the biosynthesis and processing of apoA-II, and the potential structural and processing defects in patients with dyslipopro-

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

teinemia we have cloned human apoA-II cDNA.

2. MATERIALS AND METHODS

2.1. *Preparation of a ds-cDNA library of human liver mRNA sequences*

RNA was isolated from an adult human liver obtained from a 34-year-old female accident victim (Protocol approved by a Human Research Committee) and used to establish a human liver ds-cDNA library as detailed in [15–17]. All experiments with bacteria containing recombinant plasmids were conducted in accordance with the NIH guidelines for recombinant DNA research.

2.2. *Screening the library for recombinant plasmids containing apoA-II sequences*

Three 14-base-long oligonucleotides deduced from the amino acid sequences of residues 13–17, 20–24 and 26–30 of apoA-II were purchased from OCS Labs, Denton, TX. These probes were designed to cover all possible codons of the particular amino acid sequence. For hybridization each oligonucleotide was radiolabelled at the 5′-OH end group with T_4 polynucleotide kinase (P-L Biochemicals) and [γ-^{32}P]ATP (New England Nuclear > 3000 Ci/mmol). The library was screened as in [18], except that the hybridization to the oligonucleotides was performed at 32–34°C for 40 h.

2.3. *Plasmid preparation*

Plasmids were purified by the cesium chloride density gradient procedure [19].

2.4. *Restriction enzyme analysis*

Restriction enzymes were purchased from Boehringer-Mannheim, Bethesda Research Laboratories, and New England Bio Labs. Digests were performed with the conditions specified by the suppliers.

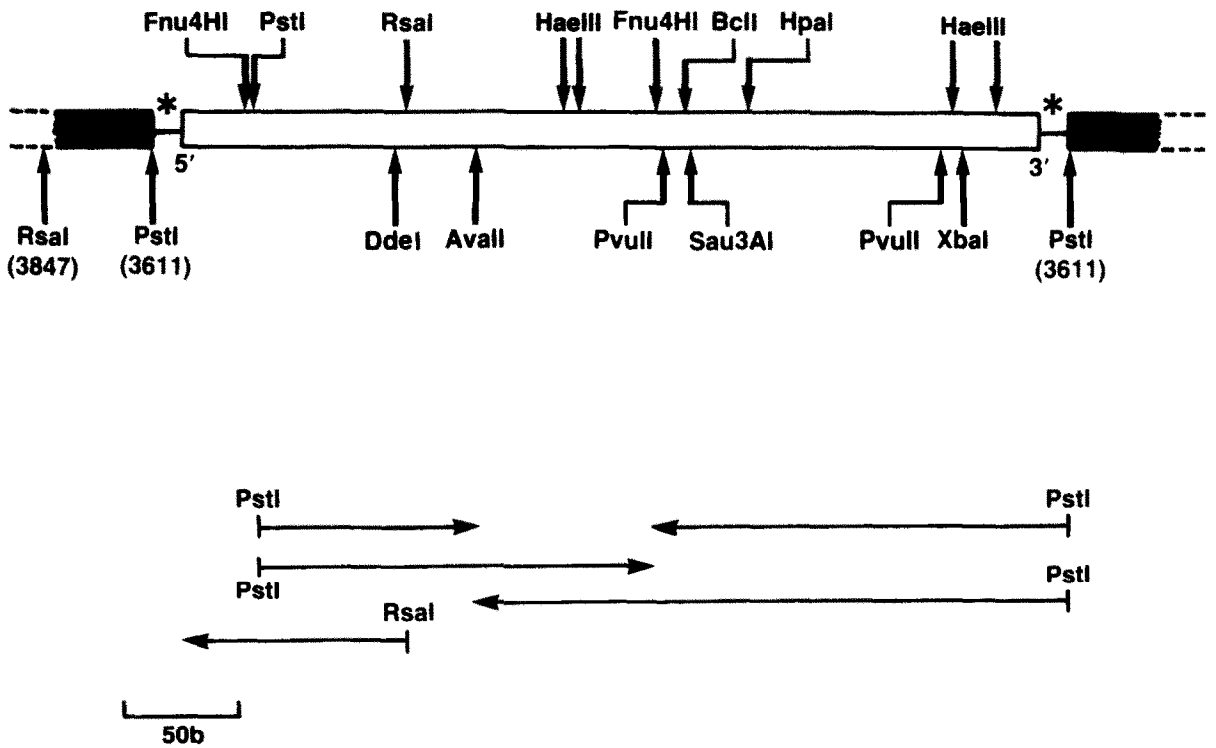

Fig.1. Restriction map of pMDB2049. Open bar, inserted ds-cDNA; solid bar, pBR322. The orientation of the coding strand within pBR322 is indicated. The asterisks indicate the dG·dC tail. The horizontal arrows indicate the length of the sequenced fragments with the arrowheads giving the direction of sequencing.

2.5. *DNA sequence analysis*

DNA fragments obtained by restriction enzyme digestion were labelled either with [α-^{32}P]ddATP (>5000 Ci/mmol, Amersham) using terminal transferase (Amersham) [20] or with [γ-^{32}P]ATP using T_4 polynucleotide kinase. DNA sequences were determined as in [21].

2.6. *5'-Extension of cDNA on liver mRNA*

To obtain a suitable primer for cDNA synthesis a restriction fragment of one clone containing the 5'-end of the insert coding for the NH_2-terminal region and part of pBR322 was isolated by agarose gel electrophoresis, labelled with ^{32}P, and digested with *Pst*I. The fragment, which served as the primer, was isolated by aqueous acrylamide gel electrophoresis, and hybridized to whole liver mRNA in 70% formamide at 37°C for 16 h [22]. cDNA synthesis was performed as in [16,17]. Primer and the extended product were sequenced together.

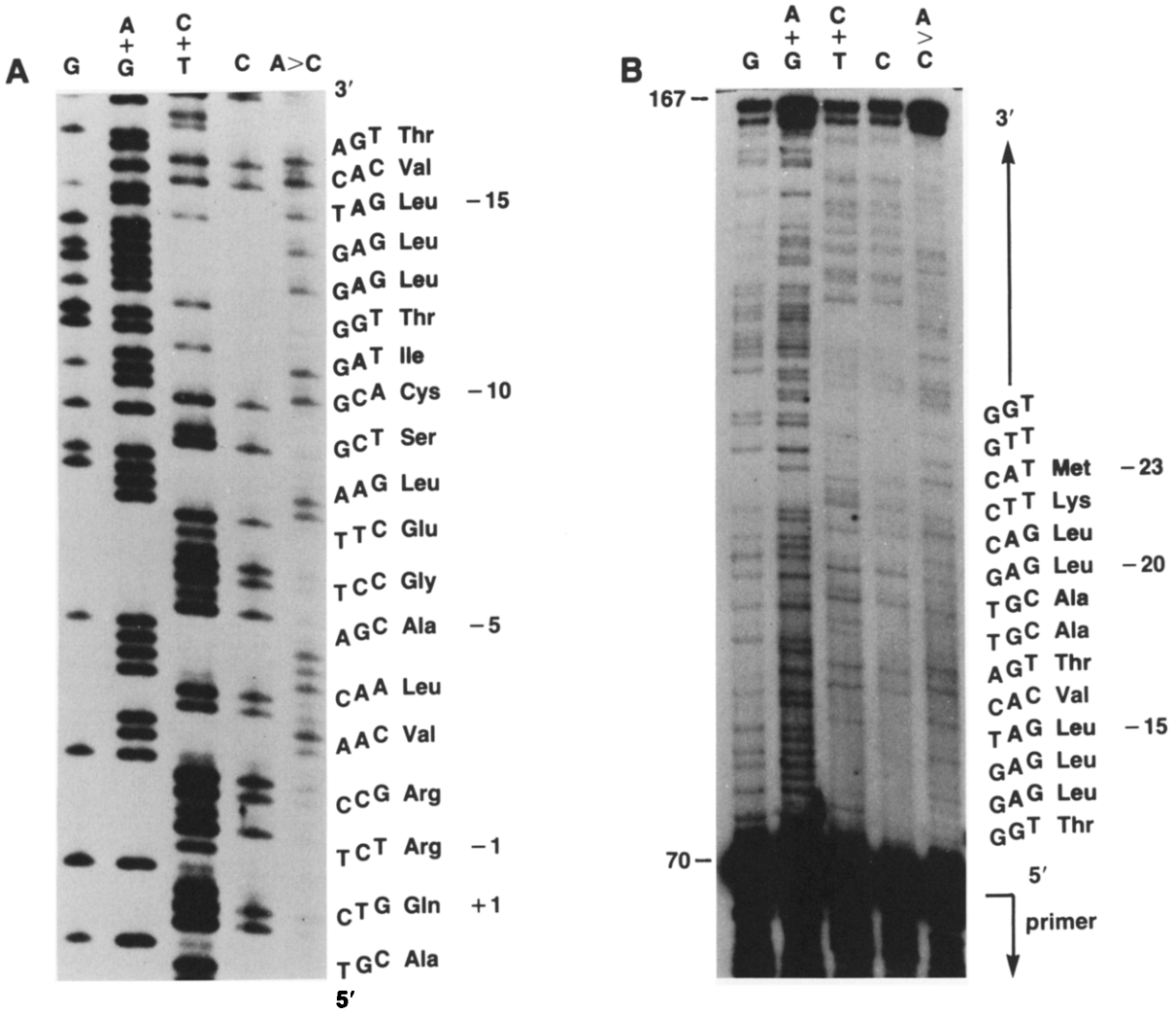

Fig.2. Maxam-Gilbert DNA sequencing ladders. (A) Sequence of the complementary strand of apoA-II mRNA encoding amino acids −17 to +2. (B) A primer extension. The 70-base-long primer ends at the codon for the Cys in position −10, which is at the bottom of the gel. The extension product at the top is 167 bases long.

3. RESULTS

3.1. *Screening for apoA-II ds-cDNA clones*

The human liver ds-cDNA library was screened for apoA-II clones with 3 different synthetic oligonucleotides (residues 13–17, 20–24 and 26–30), to maximize specificity because apoA-II contains no long amino acid sequences with little codon degeneracy. Seven out of 3200 independent clones of our ds-cDNA library hybridized to at least one of the probes, 4 of which hybridized to all 3 probes. All clones were analyzed by restriction enzyme digestion. They ranged from 350–1200 base pairs of insert length. The 4 clones which hybridized to all 3 probes were very similar, except that they differed in size, thus indicating that they were independently derived. The 3 other clones appeared to be unrelated and were not further analysed.

3.2. *ApoA-II cDNA sequence of pMDB2049*

A restriction endonuclease map of pMDB2049 is illustrated in fig.1. Based on this map restriction fragments were selected which would contain the complete nucleotide sequence of this clone. The other 3 clones had similar restriction maps except that they were shorter in length. The restriction fragments used for sequence analysis of the apoA-II clone are included in fig.1. pMDB2049 covered the protein coding region from amino acid −17 of the preprosequence to amino acid 77 of mature apoA-II. Fig.2A illustrates the complementary base sequence from amino acid −17 to +2. The other 3 clones were also sequenced using similar strategies. The base sequences common to all clones were identical.

3.3. *Complete nucleic acid sequence of human liver apoA-II mRNA*

To establish the complete nucleic acid sequence we utilized a second approach which employs primer extension of a restriction fragment of cloned apoA-II. pMDB2049 was digested with *Rsa*I which generates a blunt end fragment containing the 5′-terminal 93 nucleic acids of the coding strand, 20 bases of the dG·dC tail, and 239 nucleic acids of pBR322. This fragment was isolated, labelled at the 5′-OH ends, and cut with *Pst*I. The resulting 66/70 base fragment of pMDB2049 was hybridized to total liver mRNA, and extended with reverse transcriptase. Based on a message level of 0.2% for apoA-II mRNA [14] the molar ratio of primer to apoA-II mRNA was approx. 10:1. A single major 167 base long product was generated which indicated that pMDB2049 lacked 74 bases of the 5′-end of the apoA-II mRNA. A representative sequencing gel of a primer extension is illustrated in fig.2B. The complete nucleic acid sequence of the coding strand for preproapoA-II is shown in fig.3. The preproapoA-II mRNA contains a 58 base-long 5′-untranslated region followed by a 300 base-long protein coding region and a 3′-untranslated sequence of at least 72 bases.

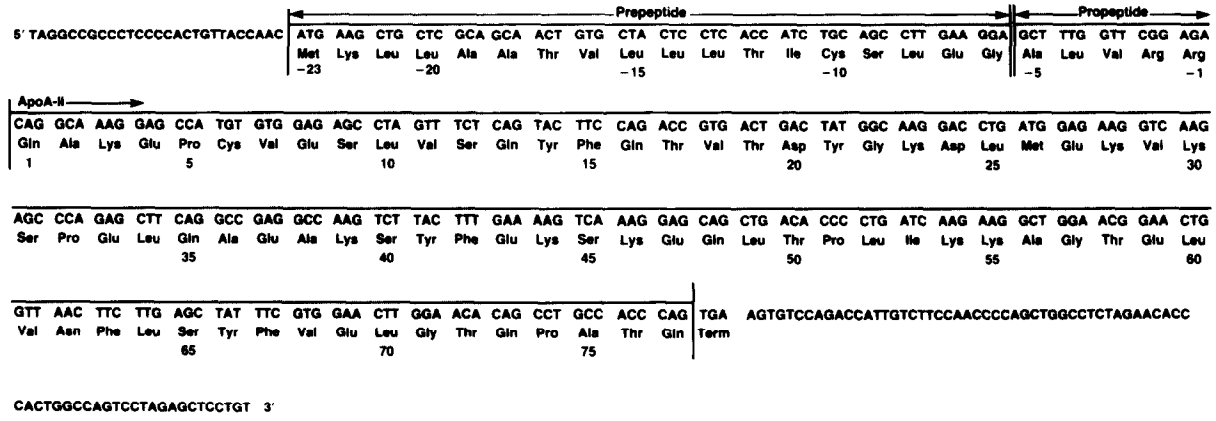

Fig.3. ApoA-II cDNA sequence and the derived amino acid sequence of preproapoA-II. The sequence from the 5′-end to the codon for amino acid −12 was established by 5′-extension sequencing. The remainder of the sequence plus an overlapping region of 17 bases was determined by sequence analysis of pMDB2049.

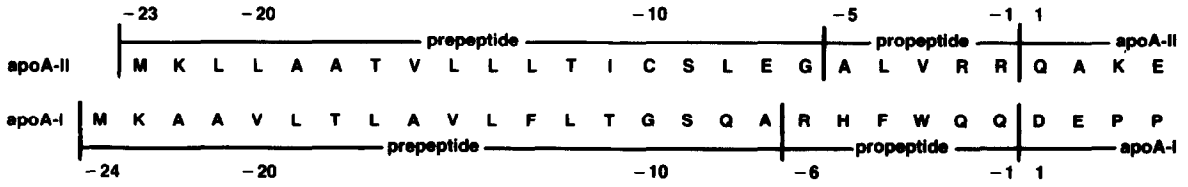

Fig.4. Comparison of preproapoA-II and preproapoA-I.

4. DISCUSSION

Here we present the complete nucleic acid sequence and the derived amino acid sequence of human liver preproapoA-II. PreproapoA-II contains an 18 amino acid prepeptide and a 5 amino acid propeptide connected to the amino-terminus of the 77 amino acid mature apoA-II. The derived amino acid sequence is in excellent agreement with the published protein sequence [10]. The only difference is a glutamic acid in position 37 instead of a glutamine. The 5'-untranslated region is most probably 58 bases long. This assumption is based on the result of the primer extension. The 3'-untranslated region is at least 72 bases long and contains a second stop-codon (TAG) in phase with the coding frame. We did not identify the typical AATAAA sequence in this region, which is probably due to incomplete second strand synthesis.

Our finding that 4 out of 3200 clones contained ds-cDNA for apoA-II is consistent with the reported level of liver apoA-II mRNA of 0.2% [14].

Of particular interest in these studies is the comparison of the biosynthesis and processing of preproapoA-II and preproapoA-I [16,17,23–25], the other major apolipoprotein of HDL. Both apolipoproteins are synthesized as preproapolipoproteins. The prepeptides do not show an extensive homology (fig.4). The propeptides are strikingly different, apoA-II having a pentapeptide terminating with two basic amino acids, while apoA-I contains a hexapeptide with a terminal gln-gln dipeptide. ApoA-I is secreted as proform, and is converted extracellularly to mature apoA-I in vitro [26–28], and in vivo in man [29]. The amino acid sequence of the cleavage site of proapoA-II, on the other hand, is similar to other secretory proproteins which usually undergo effective intracellular proteolytic cleavage [30]. The lack of complete cleavage of proapoA-II prior to secretion is unusual and is as yet unexplained.

One of the intriguing questions raised by the analysis of the biosynthesis and processing of apoA-II and apoA-I is the functional role of the proproteins. It is as yet unclear if the cleavage of the proapolipoprotein to the mature apolipoprotein is associated with the initiation or the termination of a specific physiological function. Further research will provide new and undoubtedly unexpected insights into the importance of the processing of these apolipoproteins in lipoprotein metabolism.

ACKNOWLEDGEMENT

K.J.L. is recipient of a research training grant from the Deutsche Forschungsgemeinschaft, La 499/1-1.

REFERENCES

[1] Kostner, G.M., Patsch, J.R., Sailer, S., Braunsteiner, H. and Holasek, A. (1974) Eur. J. Biochem. 45, 611–621.
[2] Scanu, A.M. and Teng, T. (1979) in: The Biochemistry of Atherosclerosis (Scanu, A.M. ed.) pp.107–122, Marcel Dekker, New York.
[3] Castelli, W.P., Doyle, J.T., Gordon, T., Hames, C.G., Hjortland, M.C., Hulley, S.B., Kagan, A. and Zukel, W.J. (1977) Circulation 55, 767–772.
[4] Avogaro, P., Bittolo Bon, G., Cazzolato, G., Quinci, G.B. and Belussi, F. (1978) Artery 4, 385–394.
[5] Fager, G., Wiklund, O., Olofsson, S.-O., Wilhelmsen, L. and Bondjers, G. (1981) Arteriosclerosis 1, 273–279.
[6] Maciejko, J.J., Holmes, D.R., Kottke, B.A., Zinsmeister, A.R., Dinh, D.M. and Mao, S.J.T. (1983) N. Engl. J. Med. 309, 385–389.
[7] Brewer, H.B. jr, Lux, S.E., Ronan, R. and John,

K.M. (1972) Proc. Natl. Acad. Sci. USA 69, 1304–1308.
[8] Edelstein, C., Noyes, C. and Scanu, A.M. (1974) FEBS Lett. 38, 166–170.
[9] Herbert, P.N., Windmueller, H.G., Bersot, T.P. and Shulman, R.S. (1974) J. Biol. Chem. 249, 5718–5724.
[10] Mahley, R.W. and Weisgraber, K.H. (1974) Circulation Res. 35, 713–721.
[11] Jahn, C.E., Osborne, J.C. jr, Schaefer, E.J. and Brewer, H.B. jr (1983) Eur. J. Biochem. 131, 25–29.
[12] Schmitz, G., Ilsemann, K., Melnik, B. and Assmann, G. (1983) J. Lipid Res. 24, 1021–1029.
[13] Stoffel, W., Krueger, E. and Deutzmann, R. (1983) Hoppe-Seyler's Z. Physiol. Chem. 364, 227–237.
[14] Gordon, J.L., Budelier, K.A., Sims, H.F., Edelstein, C., Scanu, A.M. and Strauss, A.W. (1983) J. Biol. Chem. 258, 14054–14059.
[15] Chirgwin, J.A., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[16] Law, S.W., Gray, G. and Brewer, H.B. jr (1983) Biochem. Biophys. Res. Commun. 112, 257–264.
[17] Law, S.W. and Brewer, H.B. jr (1984) Proc. Natl. Acad. Sci. USA 81, 66–70.
[18] Gergen, J.P., Stern, R.H. and Wensink, P.C. (1979) Nucleic Acids Res. 7, 2115–2136.
[19] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Handbook, Cold Spring Harbor Laboratory, New York.
[20] Yousaf, S.I., Carroll, A.R. and Clarke, B.E. (1984) Gene 27, 309–313.
[21] Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[22] Law, S.W., Tamoaki, T., Kreuzler, F. and Dugaiczyk, A. (1980) Gene 10, 53–61.
[23] Shoulders, C.C., Kornblihtt, A.R., Munro, B.S. and Baralle, F.E. (1983) Nucleic Acids Res. 11, 2827–2837.
[24] Karathanasis, S.K., Zannis, V.I. and Breslow, J.L. (1983) Proc. Natl. Acad. Sci. USA 80, 6147–6151.
[25] Cheung, P. and Chan, L. (1983) Nucleic Acids Res. 11, 3703–3710.
[26] Brewer, H.B. jr, Fairwell, T., Kay, L., Meng, M., Ronan, R., Law, S. and Light, J.A. (1983) Biochem. Biophys. Res. Commun. 113, 626–632.
[27] Edelstein, C., Gordon, J.I., Toscos, K., Sims, H.F., Strauss, A.W. and Scanu, A.M. (1983) J. Biol. Chem. 258, 11430–11433.
[28] Bojanovski, D., Gregg, R.E. and Brewer, H.B. jr (1984) J. Biol. Chem., in press.
[29] Bojanovski, D., Gregg, R.E., Ghiselli, G., Schaefer, E.J. and Brewer, H.B. jr (1984) J. Lipid Res., in press.
[30] Docherty, K. and Steiner, D.F. (1982) Annu. Rev. Physiol. 44, 625–638.