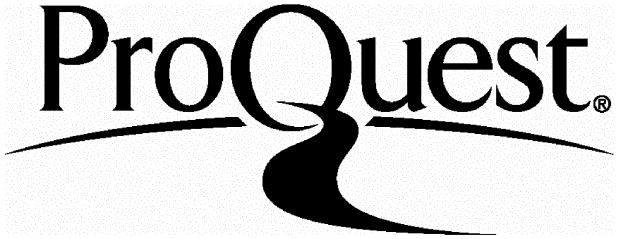
Alison Margaret Dunning

IDENTIFICATION AND CHARACTERISATION OF VARIATION IN THE HUMAN APOLIPOPROTEIN B GENE.

A thesis submitted in accordance with the regulations of the University of London
for the degree of Doctor of Philosophy
December 1993

Centre for the Genetics of Cardiovascular Disorders,
Department of Medicine,
The Rayne Institute,
University College, London.



Abstract

This thesis describes the search for genetic variants of the human apolipoprotein B (apoB) gene that have the potential to affect serum lipid levels. ApoB has a central role in lipid metabolism, both in maintaining the structure of the LDL particle and acting as the ligand for the LDL-receptor. The apoB gene is a candidate gene for the development of coronary artery disease.

Eight patients whose LDL had reduced binding to the LDL-receptor were screened for mutations in the gene region encoding the putative receptor binding domain (residues 3130 to 3630). None were found, and it was concluded that other regions of the apoB protein may also be involved in forming the ligand for the LDL-receptor. There is some evidence that regions of apoB encoded by exon 29 of the gene may also affect binding of LDL to the receptor and so this exon was also screened for variations.

A common polymorphism was discovered in exon 29 of the gene which results in the substitution of serine for asparagine at residue 4311 of the mature protein. In population samples from four different ethnic groups this polymorphism is in complete linkage disequilibrium with a reported polymorphism in exon 26 of the same gene. In addition, these two polymorphisms show complete association with the Antigen group (x/y) protein polymorphism of apoB and thus these amino acid substitutions are candidates for being the molecular bases of the Ag(x/y) epitopes. In a sample of healthy Swedish males there was a trend for the allele encoding Leu_{2712}, Ser_{4311} and Ag(x) to be associated with reduced mean serum levels of apoB and LDL-cholesterol and raised levels of HDL-cholesterol, although only the latter reached statistical significance in the population sample studied.

The high degree of linkage disequilibrium existing across the entire apoB gene enabled

a phylogenetic tree for the human apoB gene to be postulated. This parsimonious model assumed that each haplotype in the tree was derived from the preceding one by single base substitutions and one small deletion without obligate recombination events. The validity of this model was tested in a sample of Swedish myocardial infarction patients and age-matched controls. All the postulated haplotypes were seen unequivocally, however other rare, unpredicted haplotypes were also present and these occurred significantly more frequently in the patient than in the control group. Recombination events could explain the generation of the unpredicted haplotypes. Two of the unpredicted haplotypes, uniquely defined by the presence of the XbaI X- allele in combination with the allele encoding Ag(a), were seen unequivocally and were found to be associated with raised mean serum apoB levels. This reached statistical significance in the patient group.

A search was made for other amino acid substitutions encoded by exon 29 of the apoB gene in a group of individuals with aberrant binding of LDL to the receptor. Only one was found, this resulted in the substitution of threonine for arginine at residue 4243 of mature apoB.

The entire exon 29 of the human apoB gene was cloned and expressed in a baculovirus expression system. A peptide of approximately 60kDa was detected using the antibody $B_{sol}16$ on a western blot. This expressed peptide has the potential to be used to test whether the apoB region encoded by exon 29 has affinity for lipid and the LDL-receptor.

Acknowledgements

Thanks are due to everyone, too numerous to mention, who has passed through this laboratory in the last nine years for their support, friendship and for making science fun. Particular thanks are due to past and present members of the group: Rohan Taylor, Gisela Lindahl, Linda King-Underwood, David Wile, Helmut Renges, Fanli Xu, France Mailly, Rachel Peacock, Ian Day, Ros Whittle, Vilmunder Gudnason, and Anne Lane for help and stimulating discussion, to Anne Temple for keeping the laboratory running and for typing references, Simon McBride for help in sequencing the exon 29 construct, Rachel Fisher for proof-reading, Fiona Green for help with computing and to Philippa Talmud and Steve Humphries for supervision and constant encouragement.

I am also most grateful to Alan Attie and all the members of his laboratory, especially Steve Sturley, Dan Gretch, Linda Hart, Mary Kaiser and Tracey Stevens, for teaching me expression techniques and for patient, friendly help during my stay in Wisconsin.

I am indebted to our other collaborators: Johan Frostegård and Anders Hamsten for the U937 assay and the Swedish population samples, Matti Tikkanen and Réné Bütler for the Ag phenotyped Finnish population sample and to John Revill, and John Series (and their patients) for the Sheffield and Glasgow samples. Gerald Laxer, Alan Mills and Robert Brasseur gave much help with the protein modelling. Ape DNA was kindly provided by Ian Gray and Alec Jeffreys and monoclonal antibodies by Ross Milne and Linda Curtiss. The mismatch analysis in Chapter 6 was done by Helmut Renges. Thanks also to Nick Myant for his continued interest and support.

Special thanks are also due to my parents for their continued support and encouragement and to Simon Butler for everything from ironing to figure drawing and without whose support I could not have completed this course.

Financial support for this work has come through grants from the British Heart Foundation (RG5 and RG16). Travel to the USA was made possible with travel grants from the Wellcome Trust and the Nuffield Foundation.

Index

List of Tables and Figures

Abbreviations used in this Thesis

A	adenine
AcNPV	*Autographa californica* nuclear polyhedrosis virus.
Ag	Antigen group
ALP	atherogenic lipoprotein phenotype
apo	apolipoprotein
ASO	allele specific oligonucleotide
ATP	adenosine triphosphate
BMI	body mass index
C	cytosine
CAD	coronary artery disease
CETP	cholesterol ester transfer protein
Chol	cholesterol
ECL	enhanced chemiluminescence
FCHL	familial combined hyperlipidaemia
FCR	fractional catabolic rate
FDB	familial defective apoB
FDH	familial dislipidaemic hypertension
FH	familial hyperlipidaemia
G	guanine
HDL	high density lipoprotein
HL	hepatic lipase
HVR	hyper-variable region
IDL	intermediate density lipoprotein
LOD	log of the odds
Lp	lipoprotein

LPL	lipoprotein lipase
LRP	LDL-receptor related protein
mRNA	messenger RNA
MTP	microsomal triglyceride transfer protein
NMR	nuclear magnetic resonance
oligo	oligonucleotide
PCR	polymerase chain reaction
pfu	plaque forming unit
phospho.	phospholipid
RFLP	restriction fragment length polymorphism
SSCP	single strand conformational polymorphism
T	thymine
Trig.	triglyceride
VLDL	very low density lipoprotein

One and three letter codes for amino acids:-

A	Ala	alanine	R	Arg	arginine
N	Asn	asparagine	D	Asp	aspartic acid
C	Cys	cysteine	E	Glu	glutamic acid
Q	Gln	glutamine	G	Gly	glycine
H	His	histidine	I	Ile	isoleucine
L	Leu	leucine	K	Lys	lysine
M	Met	methionine	F	Phe	phenylalanine
P	Pro	proline	S	Ser	serine
T	Thr	Threonine	W	Trp	tryptophan
Y	Tyr	tyrosine	V	Val	valine

INTRODUCTION

Chapter 1

This thesis deals with the search for variant forms of apolipoprotein B (apoB) and attempts to characterise the effects of such variants on the structure and function of apoB. In order to assess the effects of variant forms of apoB, it is necessary to understand the basic biology of apoB, its role in lipid metabolism and the effects of disturbances of apoB production and clearance. Some comparisons between apoB and apolipoprotein E (apoE) will also be discussed.

1.1 Basic biology of the apolipoproteins.

The characterised apolipoproteins have three main roles: they bind lipids, stabilising and making them soluble in aqueous media; they are co-factors for enzymes and they act as ligands for lipoprotein receptors, thus providing the means by which insoluble lipids are transported and metabolised in aqueous plasma. There are specific domains within the apoproteins which provide these functions. Most of the characterised apoproteins have hydrophobic residues arranged within helices, amphipathic helices and transmembrane spanning structures which serve to bind and solubilise lipid within lipoprotein particles. The amphipathic helices, which have hydrophobic and hydrophilic faces tend to lie at the lipid/aqueous interfaces of lipoprotein particles. On apoB-100, apoE and possibly apoAI, domains consisting of charged residues exposed on the surface of the lipoproteins can interact with residues of the opposite charge on cell-surface lipoprotein receptors, enabling cells to obtain lipids from the particle. Finally structures on apoproteins which act as co-factors for enzymes, bind to the enzyme, altering its conformation and enabling the substrate to bind to the active site. An example of this is apoCII which is a co-factor for lipoprotein lipase.

1.1.1 The apoB protein.

There are two forms of apoB which serve different functions and are secreted by different tissues. The full-length protein (apoB-100) has 4536 amino acid residues with a 27/24 residue signal peptide, and the mature protein has a molecular weight of 550kDa. It is secreted by the hepatocytes (Demmer et al. 1986) as a constituent of Very Low Density Lipoprotein (VLDL). A shorter 2153 amino acid protein, termed apoB-48, is synthesised in the intestine and secreted as a part of the chylomicron. Studies using monoclonal antibodies have shown that apoB-48 is identical to the N-terminal half of apoB-100 (Olofsson et al. 1986, Hardman et al. 1987). The different forms of apoB are identified by their size on the percentile scale, with full-length apoB being apoB-100. ApoB in both its forms is secreted as lipoprotein and is not found lipid-free *in vivo*.

The amino acid sequence of apoB has very little homology with other proteins (Knott et al. 1986, Boguski et al. 1986), although it has numerous internal repeats which show homology with each other. Some functional domains of apoB have been inferred from the primary amino acid sequences. A schematic diagram of these is given in Figure 1.1.1. There are five long (52 residue) and eight shorter hydrophobic, proline-rich consensus sequences which are postulated to be lipid-binding domains, these are unique to apoB and may form amphipathic β-sheets (DeLoof et al. 1987a). There are also nine amphipathic α-helical regions, clustered particularly around the middle and at the carboxyl-terminus of the apoB-100 protein. These share consensus sequences and have hydrophobic, lipid-soluble, residues on one face and hydrophilic residues on the other and are thought to seek the lipid/aqueous interface of the lipoprotein particles (DeLoof et al. 1987b). In addition there are 39 short ($\sim$7 residue) hydrophobic sequences which are not sufficiently long to traverse a phospholipid monolayer, but are believed to anchor the apoB in the lipoprotein particle by dipping into its non-polar phase (Olofsson et al.1987). Finally, there are eight regions of basic residues, seven of which are able to bind heparin (Weisgraber & Rall, 1987).

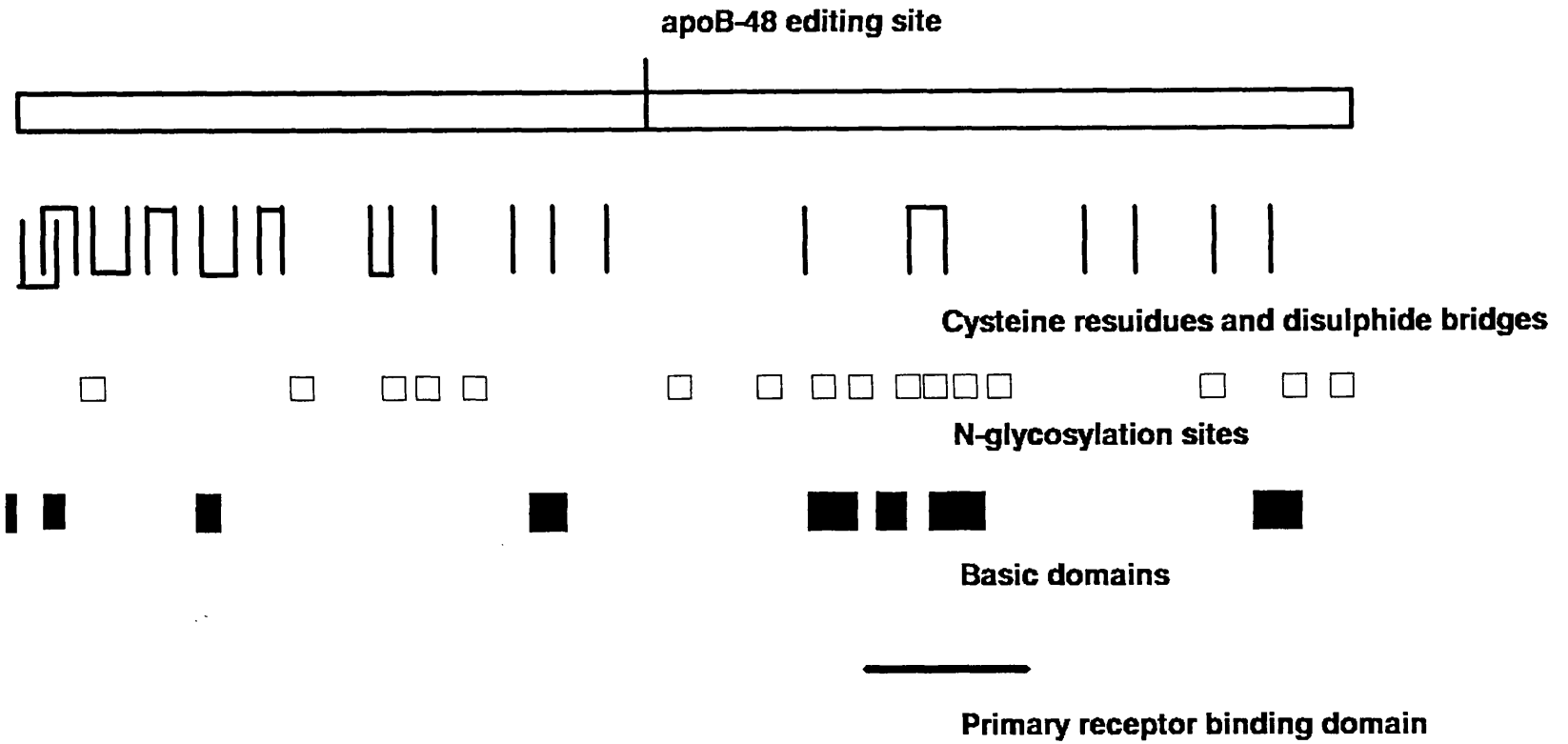
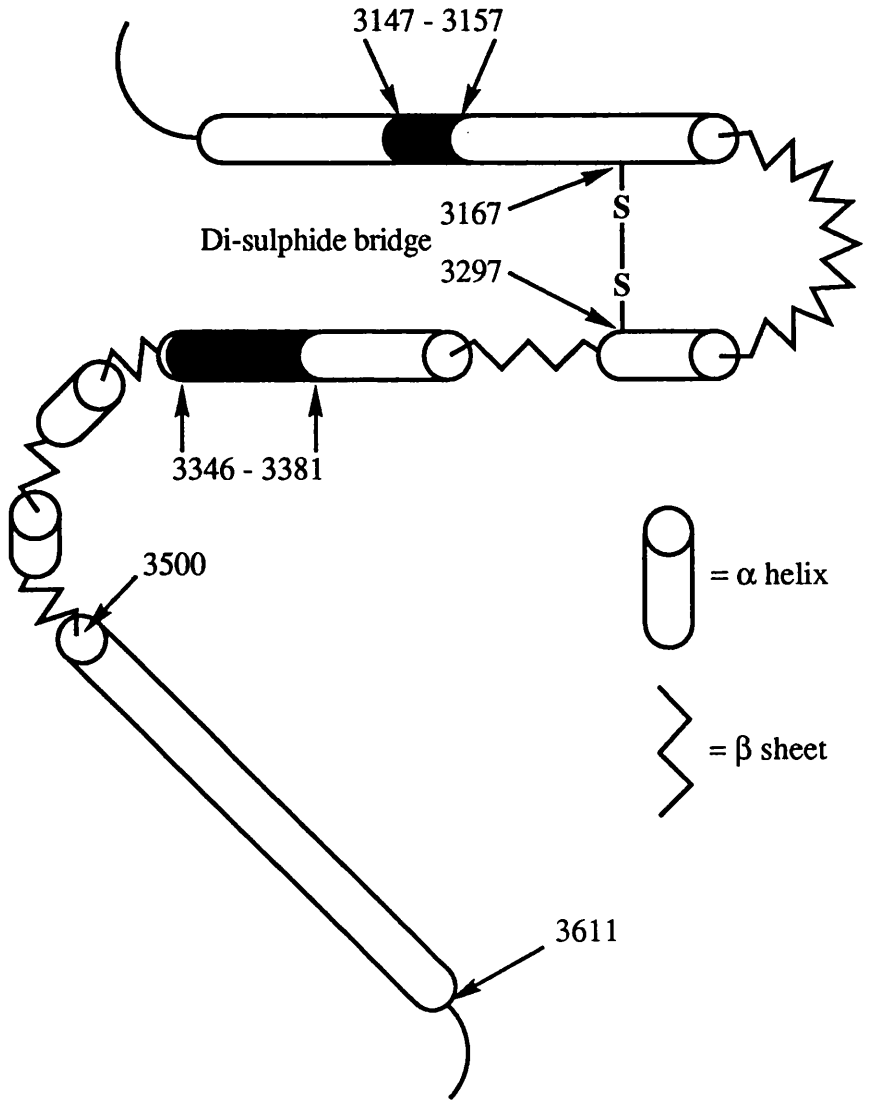


Figure 1.1.1 Map of some functional domains of apoB

(adapted from Scott, 1989 and Yang et al.1990).

Figure 1.1.2 - Representation of the primary receptor binding domain of apo B



The amino acid sequence of positively charged regions in one letter code.

3147 - 3157

```
    + +  +  +
AQYKKNKHRH
```

3346 - 3381

```
  +                +  +++   +
VIDALQYKLEGTT [RLTRKRGLK] LATALSLSNKFVE
```

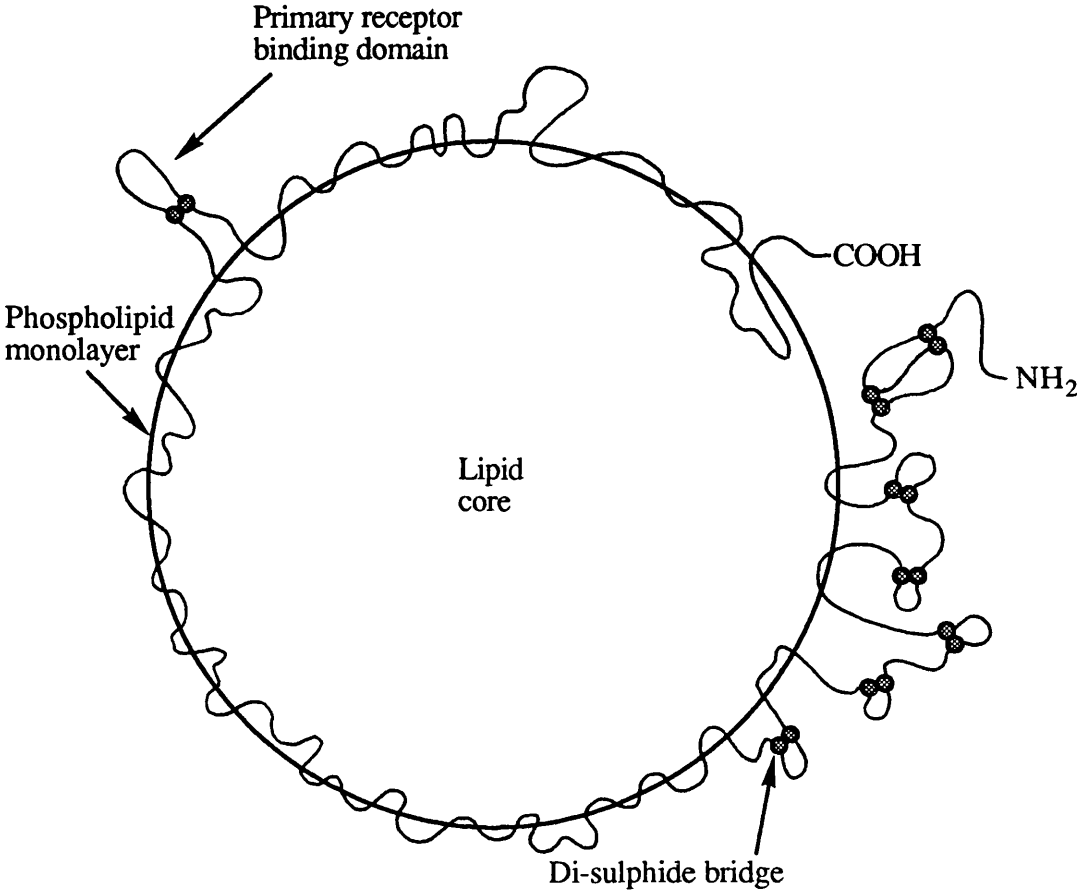
Homology to apoE

ApoB-100 has 25 cysteine residues, 16 of which participate in the formation of di-sulphide bridges (Yang et al. 1990). These are particularly clustered at the amino-terminus of the protein which consequently has a more globular structure than any other region. However, in all but two of the bridges the cross-linking is between adjacent cysteine residues, and so the effect on tertiary protein structure is not large. Sixteen asparagine residues are glycosylated (Yang et al. 1989), giving apoB a total carbohydrate content of around 10%. ApoB is phosphorylated at serine and tyrosine residues (Davis et al. 1984) and also acylated (Hoeg et al. 1986) although the functions of these two modifications are unknown.

Two of the eight regions of basic amino acids, which are particularly rich in the positively charged arginine and lysine residues, are involved in forming the primary LDL-receptor binding domain which is present on apoB-100 but not on apoB-48. A cartoon of the primary receptor-binding domain is shown in Figure 1.1.2. This region displays conservation across several mammalian species, although the net positive charge across the region does vary(Law and Scott 1990), The disulphide bridge, present in human, is not conserved in other species and so it is unlikely that this bond is crucial to the structure of the binding domain. The receptor binding domain has been defined by several methods. Firstly, there is homology between residues 3359 to 3367 of apoB and the clearly defined LDL-receptor binding domain of apoE (amino acids 142 to 150) - the other ligand of the LDL-receptor (Knott et al. 1986). These residues are predicted to form an amphipathic α-helix (Law and Scott, 1990).

ApoB(3359-3367):	Arg	Leu	Thr	Arg	Lys	Arg	Gly	Leu	Lys
	|			|	|	|		|	|
ApoE(142-150):	Arg	Lys	Leu	Arg	Lys	Arg	Leu	Leu	Arg

Figure 1.1.3 - Cartoon representation of the structure of apoB in LDL (adapted from Yang et al. 1989).



Secondly, monoclonal antibodies which bind to epitopes of apoB between residues 3000 and 3700, block the binding of apoB-LDL to the receptor (Milne et al. 1989). Thirdly, apoE-depleted lipoprotein particles carrying synthetic peptides of apoB residues 3345-3381 can bind to the LDL-receptor (Yang et al 1986).

A representation of the current idea of the structure of apoB within the LDL particle is shown in Figure 1.1.3. ApoB is woven through the phospholipid monolayer on the outside of the particle. The regions thought to be embedded within the lipid and exposed on the surface have been deduced from a combination of the postulated protein secondary structure and from digestion of LDL with trypsin to release and map the peptides exposed on the surface of the particle (Yang et al. 1989).

1.1.2 The gene for apoB

Human apoB-100 and apoB-48 are both encoded by a single gene on chromosome 2 (Knott et al. 1985, Hardman et al. 1987). The full-length gene was first cloned in 1986 (Blackhart et al. 1986, Chen et al. 1986, Cladaras et al. 1986, Law et al. 1986, Knott et al. 1986, Ludwig et al. 1987), it is 47.5kb long and has 29 exons and 28 introns (Figure 1.1.4). Introns 4, 14, 15, 20 and 21 contain Alu repeat sequences. The large size of exon 26 (7.5kb) is unusual and exon 29 (1.9kb) is also notably long. As yet, the reasons for this intron/exon structure are not understood. They could be the result of introns having been selectively lost from an ancestral gene, or due to reverse transcription from mRNA species and reinsertion into the genome. There is a poly-T sequence in intron 28 which could possibly be the residue of the poly-A tail from a reverse transcribed mRNA precursor of exon 29. The gene for apoB shares very little homology with any other genes. By contrast, the other major apoproteins: AI, AII, AIV, CI, CII, CIII and E are believed to have been derived from a common ancestral gene by duplication and they form a closely related multigene family (Li et al. 1988).

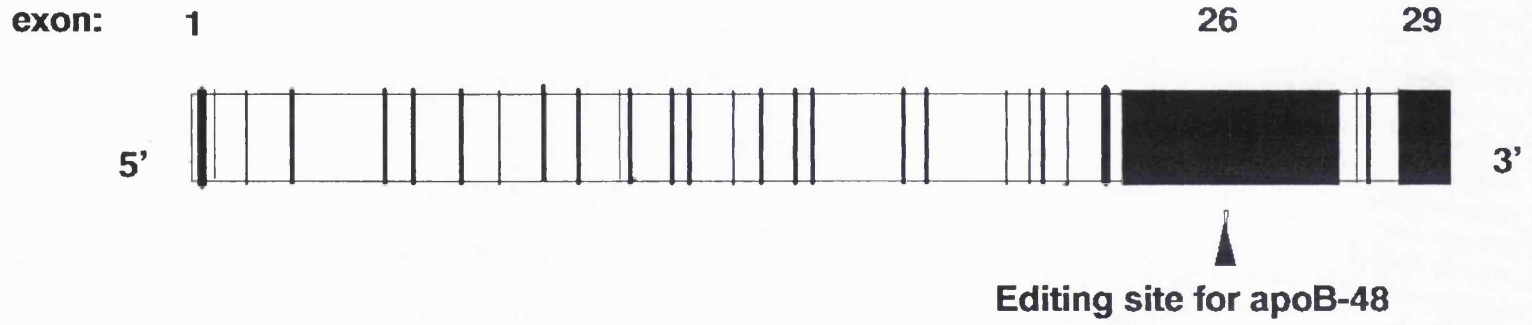


Figure 1.1.4 ApoB gene structure

Exons are shown in black (adapted from Blackhart et al.1986)

ApoB-48 is generated from the first 48% of the same gene and RNA transcript as apoB-100 (Powell et al. 1987). In human intestine an enzymic editing mechanism changes codon 2153 of the mRNA from CAA which encodes glutamine in apoB-100, to UAA - the stop codon for apoB-48. Rat intestine produces both apoB-48 and B-100 in varying ratios controlled by the amount of mRNA edited (Baum et al. 1990) and this in turn is affected by metabolic factors such as thyroid hormone (Davidson et al. 1988).

1.2 Role of apoB in lipid metabolism.

1.2.1 Lipid metabolism

A schematic diagram showing the principles of lipoprotein metabolism is shown in Figure 1.2.1. The density ranges and average compositions of the major lipoprotein classes are given in Table 1.2.2. The two forms of apoB, with their separate functions, have a central role in lipid metabolism. ApoB-48 is synthesised by the intestine and is secreted into the lymph in chylomicrons which carry triglycerides and other lipids, obtained from the diet, into the circulation. In the plasma chylomicrons acquire apoE and the apoC species by transfer from HDL. Lipoprotein lipase (LPL), on the endothelial surface of peripheral capillaries, and hepatic lipase (HL), in the liver, hydrolyse the triglycerides which are then taken up by the peripheral cells for immediate use as an energy supply or for storage in adipose tissue. LPL is activated by apoCII on chylomicrons and inhibited by apoCIII. The action of LPL generates smaller chylomicron remnant particles which are cleared by the liver, via apoE binding to the LDL-receptor related protein (LRP) (Herz et al. 1988). The exact role of apoB-48 in this pathway remains unclear; it is not a ligand for a receptor, however it's lipid-binding domains are thought to be important in maintaining the structure of the chylomicron.

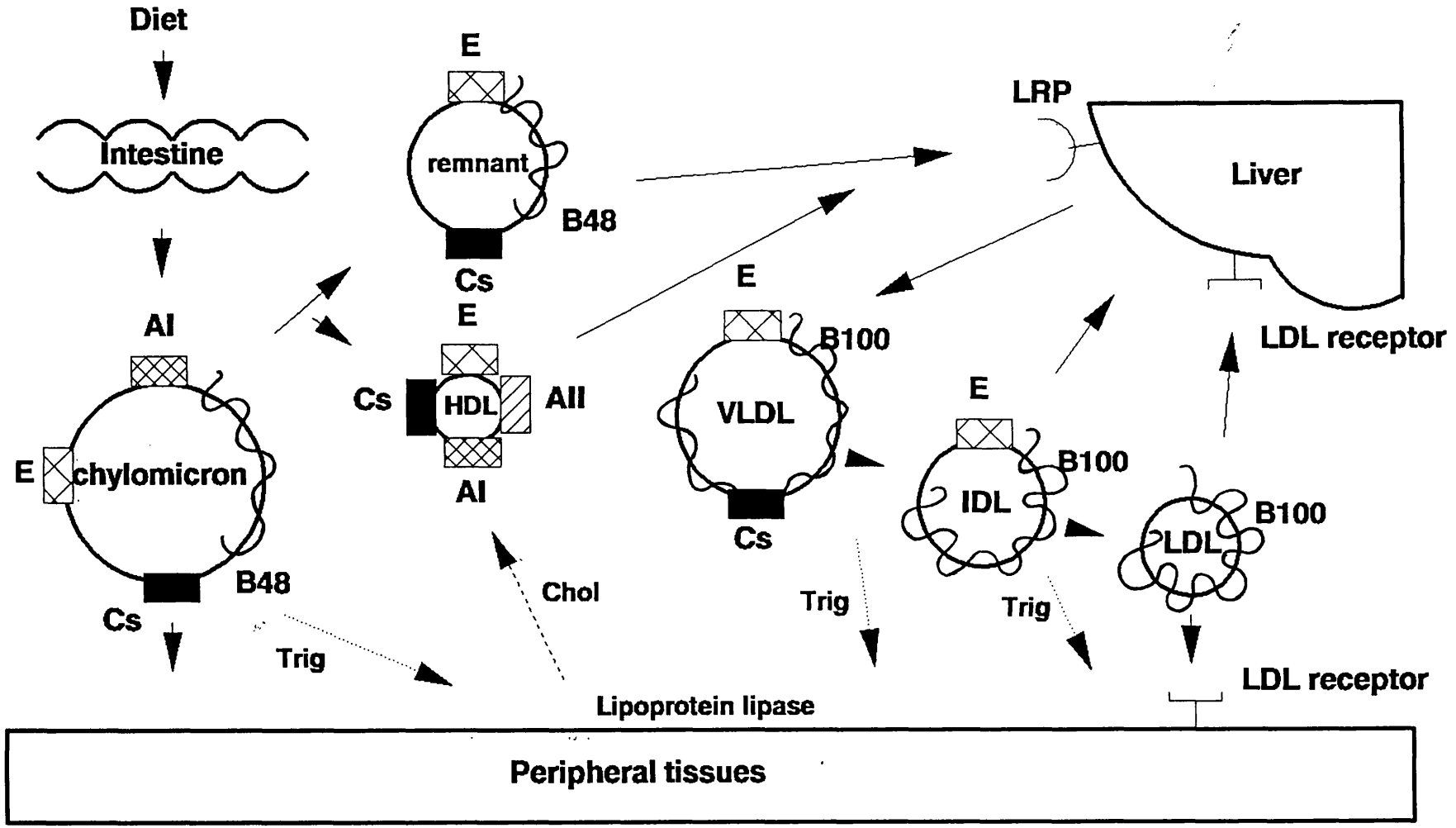


Figure 1.2.1 Schematic diagram of lipoprotein metabolism

Details in text. Figure courtesy of Fanli Xu.

Table 1.2.2 The Major Lipoprotein Classes

	Density Range g/ml	Apo-proteins	Percentage Composition				
			Protein	FC	CE	Trig.	Phospho.l
Chylo-micron	0.92-0.95	AI, AIV, B-48, CI, CII, CIII, E	2	2	4	85-90	8
VLDL	0.95-1.006	B-100, CI, CII, CIII, E	5-10	7	12	50-65	15-20
IDL	1.006-1.019	B-100, CII, CIII, E	15-20	8.5	22	20-30	20
LDL	1.019-1.063	B-100	20-25	8.5	38	7-10	15-20
HDL	1.063-1.21	AI,AII,AIV, CI,CII,CIII, E	40-55	4	11	2-5	20-35

Data from NB Myant (1990). FC = free cholesterol, CE = cholesterol ester, Trig. = triglyceride, Phospho = phospholipid

ApoB-100 is synthesised by the liver and is secreted in VLDL which transports cholesterol esters and triglycerides. VLDL additionally carries the apoproteins E, CI, CII and CIII which are also synthesised by the liver. VLDL is subject to the actions of LPL in the same manner as chylomicrons (above). Approximately 60% of VLDL-remnants are cleared directly through apoE binding to a receptor which is probably the LRP (Demant et al. 1991). The density of the remaining circulating VLDL-remnants increases, and the diameter decreases to generate first IDL and then LDL as the triglycerides are further removed by the action of LPL. IDL can

also be directly cleared via apoE binding to the LDL-receptor (Packard et al. 1985). Cholesterol ester transfer protein (CETP) also acts on the particles, catalysing the net transfer of cholesterol ester from HDL and LDL to VLDL and IDL During these processes, the other apoproteins are transferred to HDL, until apoB remains as the sole apoprotein of LDL, which is now the principal carrier of cholesterol ester. As such, apoB-100 maintains the integrity of the particle (Yang et al. 1986, 1989). In addition it serves as the ligand for the LDL-receptor, via which LDL is cleared from the circulation by the peripheral cells and principally by the liver (Brown & Goldstein, 1986).

1.2.2 Production of apoB

ApoB production and VLDL synthesis are the subject of much research and debate. Since the controversial aspects are probably not relevant to the subject of this thesis, they will not be discussed here.

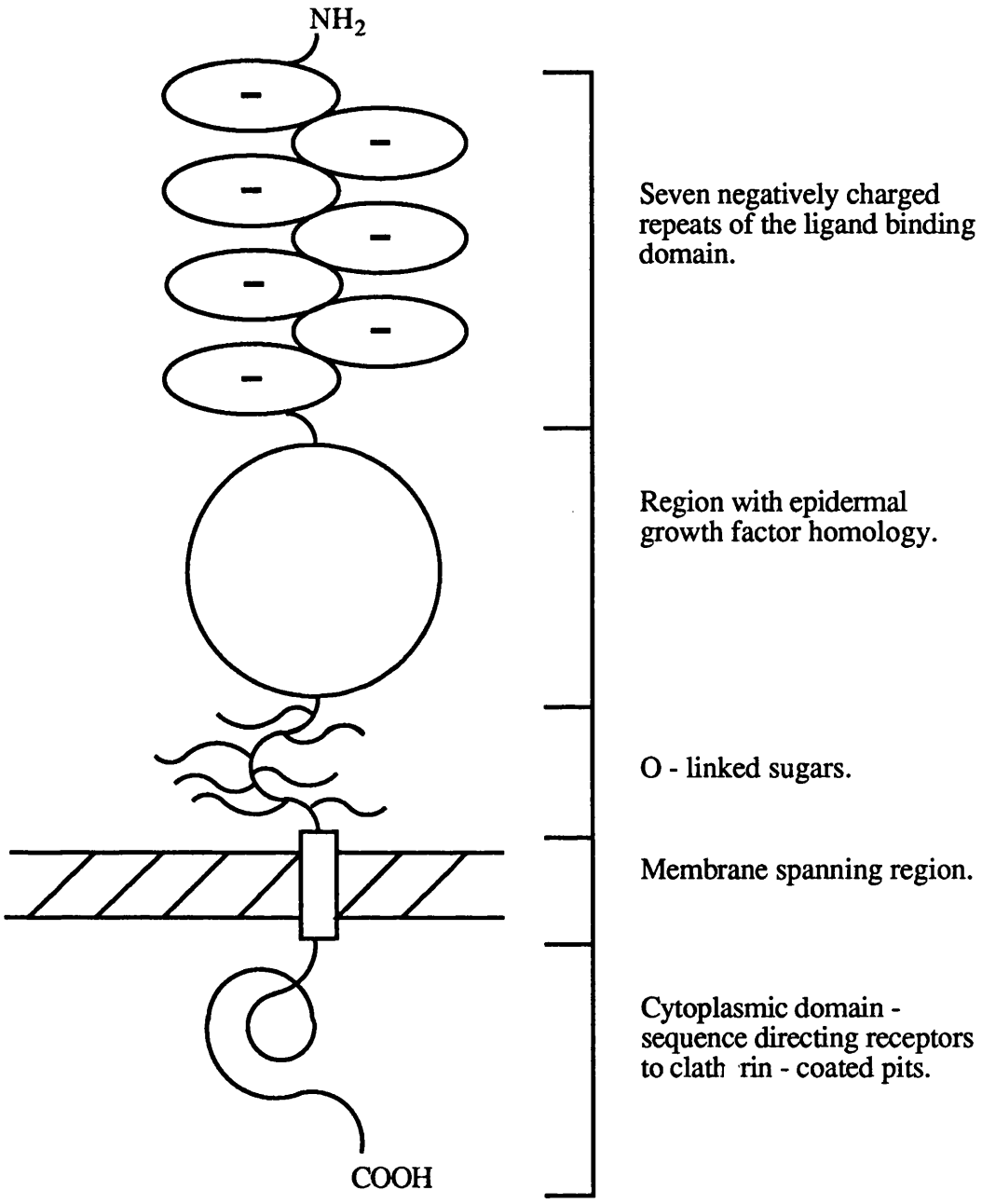
The apoB gene is transcribed constitutively and spliced to form a mRNA of 14kb in both liver and intestine. Transcription does not appear to be influenced by metabolic factors and the mRNA half-life is more than 16 hours in HepG2 cells (Pullinger et al. 1989) indicating that RNA stability is not a controlling factor in apoB protein synthesis. The mRNA takes approximately ten minutes to be translated on the ribosome (Borchardt and Davis 1987, Bostrom et al. 1988) and the nascent protein is co-translationally attached to the inner membrane of the endoplasmic reticulum (Knott et al. 1986, Olofsson et al. 1987). On the inner membrane or in the lumen of the endoplasmic reticulum apoB becomes associated with phospholipids, triglycerides, unesterified cholesterol and cholesterol ester. If there is insufficient lipid available, the apoB is degraded within the endoplasmic reticulum or possibly the Golgi apparatus (Borchardt and Davis 1987, White et al. 1992). If sufficient lipid is available, the apoB, probably attached to the phospholipid monolayer derived from the endoplasmic reticulum and surrounding a core of lipids,

buds-off to form a nascent lipoprotein (Olofsson et al. 1987). This passes through the smooth endoplasmic reticulum and into the Golgi apparatus where more lipid is added (Janero and Lane 1983, Higgins 1988) before the mature lipoprotein is secreted. In HepG2 cells apoB secretion is increased by the addition of oleic acid and decreased by fish oils, insulin and albumen (Dashti et al. 1989, Pullinger et al. 1989). Regulation of apoB production is at the level of lipoprotein secretion and is mediated via apoB degradation, although it is not known if degradation is a an active control process or a default mechanism. Control of apoB production is however post-translational and not transcriptional. This method of control of apoB production ensures that there is always sufficient intestinal and hepatic apoB available to efficiently package all dietary lipid input and hepatic lipid output as chylomicrons and VLDL.

1.2.3 Clearance of apoB - interaction with the LDL-receptor.

As stated in section 1.2.1, all of apoB-48, as chylomicron remnants, and a proportion of apoB-100, as VLDL-remnants, are cleared via the binding of apoE to the LRP and LDL-receptor. This clearance, which follows the hydrolysis of triglycerides from the particle by LPL, is rapid - taking only a few minutes for chylomicrons to be cleared from the circulation. The receptor binding domain of apoB only becomes available to the LDL-receptor once the VLDL particle has been metabolised to LDL (Chappell et al. 1991). On VLDL and IDL either apoB is in an altered conformation, or it is not exposed on the surface of the particle, since it is not recognised either by the LDL-receptor or by monoclonal antibodies directed towards it (Krul et al. 1985). However, since the majority of cholesterol ester in the circulation is transported as LDL, the role of apoB as the ligand through which LDL is cleared is crucial to cholesterol metabolism. Clearance of LDL takes 36 hours to three days. The positively charged residues of apoB are thought to interact with the negatively charged aspartate and glutamate amino acids in the seven ligand-binding domains of the LDL-receptor (Brown and Goldstein, 1986) (Figure 1.2.3).

Figure 1.2.3 - Cartoon of the domains of the LDL - receptor



Consensus sequence of the seven ligand binding domains (Goldstein et al, 1985)

Cys – [Asp] – X – X – X – [Asp] – Cys – X – [Asp] – Gly – Ser – [Asp] – [Glu]

Lund-Katz and colleagues (1988) used ^{13}C NMR to investigate apoB lysine residues, 63% of which, are exposed on the surface of LDL and found that they exist in two chemical microenvironments, dependant upon the local conformation of the apoB. About 170 of the exposed Lys residues have an ϵ-amino group with a pK of 10.5, whilst 53 "active" lysines in basic or non-polar microenvironments have a pK of 8.9. These pK 8.9 Lys have a lower affinity for protons than the others and it is energetically favourable for them to form electrostatic interactions with the negatively charged groups such as those on the LDL receptor or heparin (Section 1.2.4). There are 10 active lysines in very basic microenvironments in the C-terminal portion of apoB, with six of these being in the primary receptor binding domain, and it is most likely that these are responsible for interaction with the receptor.

The seven ligand-binding domains of the LDL-receptor are each capable of binding one molecule of apoE but co-operate to bind apoB. Thus, the LDL-receptor can bind four to seven apoE molecules simultaneously, dependant upon their spacial arrangement on the VLDL particle, but only a single apoB molecule. There is evidence that apoE binds to the LDL-receptor as dimers (Dyer et al. 1990). This is reflected in the number of apoproteins per lipoprotein particle; VLDL have four to six apoE proteins whilst both VLDL and LDL have only one molecule of apoB. (Russell et al. 1989). The relative sizes of the receptor and lipoprotein ligands also affect their interaction. Most LDL-receptors are clustered within clathrin-coated pits which occupy less than 2% of the cell surface, the LDL particle is 20 times larger than the receptor and electron-microscopy studies have revealed many LDL particles clustered within an individual pit. Chappell et al. (1991) have shown in *in vitro* studies that crowding of receptor-bound particles sterically hinder further particles from binding. These studies demonstrated that LDL-receptors can clear fewer large than small LDL particles because, when bound to a receptor, the larger particle blocks access of other particles to adjacent LDL-receptors within the pit. Additionally

these studies revealed that large LDL particles, which still carry apoE, can simultaneously bind to two LDL-receptors - using apoB to bind to one and apoE the other. The size differential means that a single LDL-receptor can only bind one lipoprotein particle, even though it has seven sites for binding apoE. The presence of multiple apoEs on a VLDL particle probably enhance its clearance by increasing the number of orientations in which it can interact with the receptor, whilst LDL-apoB can only interact in one orientation.

1.2.4 Interaction of apoB with heparin and arterial-wall proteoglycans.

ApoB containing lipoproteins can interact with proteoglycans such as heparan sulphate and chondroitin sulphate found on the arterial wall. There are seven regions of basic amino acids in apoB that can bind heparin (Section 1.1.1, Weisgraber & Rall, 1987) and at least three that can interact with chondroitin sulphate-rich aortic proteoglycans (Camejo et al. 1988). The primary receptor binding domain of apoB coincides with regions identified in both the above studies, which may account for the fact that heparin is capable of displacing LDL from the LDL-receptor (Goldstein et al. 1976). A physiological purpose of this binding could be to enable triglyceride rich lipoproteins to bind to the endothelial surface prior to hydrolysis by lipoprotein lipase (Zilversmit, 1973) which is also anchored to the surface of endothelial cells by heparan sulphate (Cheng et al. 1981). However, the fact that triglyceride-rich VLDL have a lower affinity for arterial proteoglycans than LDL (Camejo et al. 1988) does not support this hypothesis, and may indicate that these regions only become exposed after hydrolysis of the triglyceride from the particle. The interaction is believed to be mediated by electrostatic forces between positively charged residues (Arg and Lys) on apoB and the negative charges on heparin. This interaction, which binds LDL to the endothelial surface of arteries and capillaries, may be a contributing factor in atherogenesis (Hollander, 1976).

1.2.5 The Lp(a) particle

The ability of apoB to bind to the receptor can also be reduced by its covalent attachment to apo(a) to form Lp(a) particles. First identified by Berg (1963), high serum Lp(a) levels are an independent risk factor for the development of CAD (Kostner et al. 1976, Armstrong et al. 1986). Lp(a) particles are larger and more dense than LDL since apo(a) has a molecular weight of between 400 and 700kDa. The amino acid sequence of apo(a) was deduced by McLean et al. (1987), and shows striking homology with plasminogen. Plasminogen is the zymogen of plasmin, a protease which digests fibrin clots, it has five homologous structures called kringles (due to their resemblance in shape to a Danish biscuit), and a C-terminal protease domain. The gene for apo(a) has been derived from the plasminogen gene by duplication, and the two genes lie within a 50kb region of chromosome 6 (Malgaretti et al. 1992). The apo(a) gene has its own promoter (Wade et al. 1993), it has lost the N-terminal domain and kringles 1 to 3, but does encode many repeats of kringle 4, followed by one kringle 5 and an inactive protease domain. Lp(a) particles vary markedly between individuals in both size and plasma levels. The size variation is due to the different numbers of repeats of kringle 4, which can range from 2 to more than 20 (Lackner et al. 1991). The serum levels of apo(a) are inversely related to the size (Utermann et al. 1987). More than 90% of variation in Lp(a) levels are due to differences in the apo(a) gene (Boerwinkle et al. 1992), although the genes for apoE and the LDL-receptor also have an effect on Lp(a) levels (DeKnijff et al. 1991, Leitersdorf et al. 1991). Apo(a), like apoB, is synthesised in the hepatocytes but it is not yet clear where or when during production the two become covalently linked. Studies in baboons suggest that linkage could occur in the medium surrounding the hepatocytes or just before secretion from the cells (White et al. 1993) and studies with transgenic mice expressing apo(a) and infused with human LDL indicate that linkage can occur in the plasma (Chiesa et al. 1992). The disulphide bridge that forms this covalent link is believed to involve Cys_{3734} or Cys_{4190} in apoB (Coleman et al. 1990) and a Cys residue in the penultimate repeat of kringle 4 in apo(a) (McLean et al. 1987). When apoB and apo(a) are covalently linked, the

lipoprotein produced is not VLDL but Lp(a). Lp(a) has low affinity for the LDL-receptor - individuals with homozygous and heterozygous LDL-receptor defects catabolise Lp(a) at the same rate as normal individuals (Krempler et al. 1982, Knight et al. 1991). Furthermore, individuals with the apoB Gln_{3500} mutation (Section 1.4.3) accumulate Gln_{3500}-LDL, but not Arg_{3500}-Lp(a), in their plasma (Perombelon et al. 1992). Armstrong et al. (1986) showed that skin fibroblasts *in vitro* degrade Lp(a) slower than LDL. However, the reason why attachment of apo(a) to apoB affects the receptor-binding domain of apoB is not understood. The alternative mechanism, through which most Lp(a) particles are catabolised, is not elucidated either. Native Lp(a) are not significantly taken up by the macrophage scavenger receptor (Snyder et al. 1992). However, Lp(a) particles are more susceptible to oxidation than LDL (Naruszewicz et al. 1992a & 1992b), and once oxidised they are cleared by the scavenger receptor (Haberland et al. 1992). Williams et al. (1992a) have also demonstrated that fibroblasts take up native Lp(a) in the presence of LPL, this does not involve the lipolytic function of LPL but may use the LPL associated heparan sulphate proteoglycans. If demonstrated to have a physiological role this mechanism could also lead to atherogenesis.

In addition to poor clearance of Lp(a) leading to high serum cholesterol levels and other lipid-related atherogenic factors, Lp(a) may also have an adverse effect on the clotting cascade. Lerch et al. (1980) have demonstrated that kringle 4 of Lp(a) has weak binding to fibrin but no protease catalytic activity. Harpel et al. (1989) showed that Lp(a) competes with plasminogen for binding to fibrin and fibrinogen, and thus may inhibit the normal degradation of blood clots. In addition, if Lp(a) binds to the fibrin which is present at the site of arterial wall injury, it may supply the regenerating artery with a supply of cholesterol. In moderation, this would be important for the synthesis of new cell membranes, but in excess it may initiate atherogenesis. Rath et al. (1989) and Cushing et al. (1989) have demonstrated the presence of apo(a) in human atherosclerotic plaques but not in normal aorta. Furthermore, transgenic mice expressing human

apo(a), on a high fat diet, develop atherosclerotic lesions which also contain apo(a) (Lawn et al. 1992), indicating that Lp(a) does play a direct role in atherogenesis.

1.2.6 Oxidised LDL

LDL is subject to free-radical damage and oxidation which particularly alters the positively charged lysine residues on apoB (Steinbrecker et al. 1989) and abolishes its ability to bind to the LDL-receptor. Ultimately oxidation leads to the fragmentation of apoB within the particle (Schuh et al. 1978). *In vitro* oxidation can also result in the apparent aggregation of LDL particles (Gandjini et al. 1991). The mechanisms by which LDL are oxidised *in vivo* are not fully understood, but the reaction may be mediated by cells within the arterial wall. LDL are oxidised *in vitro* by incubation with macrophages, endothelial cells and smooth muscle cells (Parthasarathy & Steinberg 1992). In addition, phospholipase A2 has a role in the modification of LDL associated with oxidation, and there is evidence that apoB has an intrinsic phospholipase A2 activity (Parthasarthy & Barnett, 1990), indicating that apoB may induce some of the damage itself. LDL which are cleared slowly tend to undergo more damage since increased plasma levels of LDL encourage influx of the particles into the arterial wall and the anti-oxidants, such as vitamin E and β-carotene present within the lipid, get depleted with time. In addition, small dense LDL particles (Section 1.3.3) and LDL with reduced unesterified cholesterol content have increased susceptibility to oxidation (Tribble et al. 1992), perhaps because the surface conformation of such particles makes susceptible regions of the particle more accessible to oxidising agents. There is evidence that hydrophobic epitopes of apoB associated with lipid are more susceptible to oxidation than those exposed on the outside of the lipoprotein (Gandjini et al. 1991). This may be due to peroxide radicals, generated within the lipid, directly affecting the adjacent apoB residues. Such oxidation-damaged particles are cleared by the macrophage scavenger receptor (Steinberg et al. 1989). Monocytes/macrophages which become cholesterol-filled through scavenging LDL are called foam-cells and can initiate atherogenesis (Section 1.4.1).

1.2.7 The role of apoE in lipid metabolism.

ApoE is the other ligand for the LDL-receptor, although it has homology to apoB only in the short receptor-binding region. Mature apoE is 299 residues in length and is secreted primarily by the hepatocytes (Mahley et al 1988). Unlike apoB, it is able to move between lipoproteins and there are normally several apoE molecules per lipoprotein particle - it is now believed that apoE may function as a dimer (Dyer et al. 1990). The tertiary structure of apoE is better understood than that of apoB. It consists of two, mostly α-helical domains connected by a hinge region (residues 165-200) (Wetterau et al. 1988). The structure of the amino-terminal domain has been determined by X-ray crystallography (Wilson et al. 1991), its main structural features are four long (28-36 residue) α-helices arranged in a bundle with the receptor binding domain (residues 134-160) on the fourth helix. The receptor binding domain has also been defined by examining the amino acid substitutions in mutant forms of apoE which do not bind to the receptor (Table 1, Rall & Mahley 1992). The carboxyl-terminal domain of apoE predominantly consists of lipid binding amphipathic α-helices. In the absence of lipid this domain mediates the self-association of apoE into tetramers (Aggerbeck et al. 1988).

There are three common isoforms of apoE (Rall et al. 1982), which result from charge differences at residues 112 and 158, and these can be detected by isoelectric focusing. The amino acid differences generating these variant forms are shown in Table 1.2.7. Other rare variants of apoE may have the same charge as these but different amino acid substitutions and these will not be dealt with here. The frequencies of the common apoE variants differ between ethnic groups (Davignon et al. 1988), however the overall high frequency of the apoE3 form make it the likely wild-type.

	Residue 112	Residue 158	Relative charge	Allele Frequency*
Apo E2	Cys	Cys	0	0.073
Apo E3	Cys	Arg	+1	0.783
Apo E4	Arg	Arg	+2	0.143

*Overall frequency (Davignon et al. 1988, Table 1)

Table 1.2.7 The amino acid changes that generate the common apoE variants.

The arginines at residues 112 and 158, cause apoE4 to have more positive charges than apoE3, which in turn is more charged than apoE2. ApoE2, lacking the arginine residues, has low affinity for both the LDL-receptor and LRP (Mahley et al. 1984, Kowal et al. 1990) and lipoproteins containing apoE2 are catabolised slowest of the three isoforms (Gregg et al. 1981). By contrast, apoE4 containing lipoproteins are catabolised fastest (Gregg et al. 1986) although their affinity for the receptor is the same as that of apoE3.

Association studies on normolipidaemic individuals with the different isoforms have demonstrated that there is a ranked effect of apoE genotype on mean cholesterol levels; with individuals homozygous for the E2 allele having the lowest levels, through intermediate levels in individuals with apoE3 alleles, up to the highest levels in individuals carrying one or more E4 alleles (Utermann et al. 1979b, Bouthillier et al. 1983) The same association is seen between apoE genotype and mean apoB levels (Utermann et al. 1985, Sing and Davignon, 1985). However apoE genotype has the opposite effect on mean levels of apoE - the apoE2 allele is associated with the highest mean level of apoE and the E4 allele with the lowest (Utermann et al.

1985). When the average effects of the apoE alleles, on mean total cholesterol levels are considered in seven population studies (Davignon et al. 1988), the apoE2 allele lowers mean cholesterol levels by 6 - 9%, and the apoE4 allele raises them by 3 - 4.5%. Thus, although apoE2 is catabolised slowest, it is associated with the lowest levels of cholesterol and apoB, whilst the rapidly cleared apoE4 allele is associated with the highest levels.

The reason for this association is that apoE modulates the rates of both production and clearance of apoB containing lipoproteins (Davignon et al. 1988). ApoE2 containing lipoproteins have a lower affinity for the receptors, and are catabolised via apoE more slowly than apoE3 or E4 containing lipoproteins. This results in accumulation of apoE2 containing chylomicron- and VLDL-remnants in the plasma. In addition apoE2-VLDL and IDL are not good substrates for LPL and are not easily hydrolysed to LDL (Ehnholm et al. 1984b), thus reducing the levels of LDL in the plasma. The low rate of delivery of cholesterol to the liver, due to slow clearance of the chylomicron- and VLDL-remnant particles, causes the liver to up-regulate the number of LDL-receptors on its surface, which in turn increases the clearance of LDL using apoB as the ligand. In apoE2 homozygotes the up-regulation of the LDL-receptor and lack of competition from apoE possibly also enable small VLDL and IDL to be catabolised directly via apoB binding to the LDL-receptor (Demant et al. 1991). The net result of all this is decreased production and increased removal of LDL, hence the association of apoE2 with lower levels of apoB and total cholesterol. However, the slow clearance of apoE2-VLDL and chylomicrons does lead to increased plasma triglyceride concentrations which will be discussed in section 1.2.8. ApoE4 has an opposite effect on plasma cholesterol and apoB levels to apoE2, it has an affinity for triglyceride-rich lipoproteins (chylomicrons and VLDL) whilst apoE3 and E2 preferentially associate with HDL (Gregg et al. 1986, Steinmetz et al. 1989). Weisgraber (1990) has elucidated the mechanism underlying this: The Cys_{112} residue in apoE3 enables it to form dimers with apoAII and so it preferentially migrates to HDL. ApoE4, which has Arg_{112}, is not able to bond

to apoAII and so remains on VLDL. As a result, the VLDL and chylomicron remnants of individuals with an apoE4 allele are likely to carry more apoE molecules. Lipoproteins with more apoE are cleared rapidly via this apoE, delivering cholesterol and triglycerides to the liver, which responds by down-regulating the LDL-receptor. This generates a decreased removal rate of LDL via apoB, leading to higher total cholesterol and apoB levels, but unchanged triglyceride levels.

1.2.8 Interaction of apoE genotype with other factors to cause disease.

A different effect of apoE genotype is seen in individuals with Type III hyperlipoproteinaemia. Type III hyperlipoproteinaemia is defined as the plasma accumulation of chylomicron- and VLDL- remnants in patients who have elevated plasma concentrations of both cholesterol (8 - 20mM) and triglycerides (3.5 - 8mM) and develop xanthomas of the palmar creases and tuberous xanthomas (Mahley & Rall,1989). The frequency is estimated to be between 4/10 000 and 1/100 000 in the population (Mahley and Rall,1992). These patients have a high incidence of premature vascular disease. Approximately 95% of type III hyperlipoproteinaemia is associated with apoE2 homozygosity (Utermann et al. 1979a) but it does have other causes including apoE deficiency (Schaefer et al. 1986, Cladaras et al. 1987). Although individuals with the E2/E2 genotype are very frequent among groups of type III patients, the converse is not true - less than 5% of individuals homozygous for the apoE2 allele develop type III hyperlipidaemia (Davignon et al. 1988). The explanation for this effect is that while the slower metabolism of apoE2 particles normally leads to decreased serum cholesterol levels and slightly increased triglycerides (section 1.2.7), all E2 homozygotes have disbetalipoproteinaemia (slow clearance of chylomicron- and VLDL-remnants). In about 5% of these individuals, secondary genetic or environmental risk factors cause them to develop type III hyperlipidaemia (Mahley & Rall 1989). The other risk factors are: an often postulated second genetic lipid abnormality; hypothyroidism, diabetes and other hormonal factors; diet; drugs and increasing age. In addition

to apoE2 homozygosity, some rare variants of apoE which are also defective in receptor binding, have been found to cause type III hyperlipidaemia. Unlike apoE2, however, some of these rare variants generate dominantly inherited disease (Rall & Mahley 1992). The aetiology of type III hyperlipoproteinaemia clearly illustrates the multifactorial and polygenic nature common to many of the hyperlipidaemias.

Recently, the apoE4 isoform has been associated with Alzheimer's disease, another multifactorial disease of older-adults, characterised by brain lesions, and progressive dementia until death. Corder et al. (1993) have studied 42 families with late-onset Alzheimer's disease and found an association of the apoE4 allele with incidence and age-of-onset of the disease in an allele dose-dependent manner. Whilst only 20% of family members with no apoE4 isoform were affected, 47% of those with one apoE4 allele and 91% of those homozygous for the apoE4 allele were affected, and age-of-onset of symptoms decreased with increasing apoE4 allele-dosage. To date, no population association studies have been performed to ascertain the risk of developing Alzheimer's disease for individuals, in the general population, with the apoE4 allele - it is possible that only a subset of individuals are at risk, as with E2 homozygotes and type III hyperlipidaemia. The association could be due to linkage disequilibrium with alleles of other genes situated close to the apoE gene on chromosome 19. However, *in vitro* studies have shown that apoE4 has a greater affinity for β-amyloid protein than the other isoforms, perhaps due to interaction with apoE4's increased positive charges. Mutation in the gene for β-amyloid has already been demonstrated to be one cause of early-onset Alzheimer's disease (Goate et al. 1991), and more β-amyloid protein is seen deposited in the brain lesions of homozygous apoE4 Alzheimer's patients than in others. Other studies have demonstrated that both apoE and apoB are found the brain lesions of Alzheimer's patients, although apoB is not found in healthy brain tissue (Namba et al. 1991, 1992). It will be interesting to see if Alzheimer's disease becomes reclassified as a disorder of lipid-metabolism as a result of this and future studies.

1.3 Disorders of ApoB production.

Two rare genetic apoB deficiency disorders have been characterised: abetalipoproteinaemia and hypobetalipoproteinaemia (Kane and Havel 1989). A collection of disorders with the common characteristic of over-production of apoB, termed familial combined hyperlipidaemia (FCHL) have also been described, and studies to elucidate their exact molecular bases are still on-going (Kwitero vich et al. 1993). The study of these apoB production disorders has revealed more about the role of apoB in lipid and lipoprotein metabolism.

Both apoB deficiency disorders are characterised by failure of the liver and intestine to manufacture apoB containing lipoproteins: VLDL, LDL and chylomicrons. The main problems associated with these diseases result from malabsorption of dietary fat and fat-soluble vitamins. The vitamin deficiencies, if untreated, cause neural damage such as spinocerebellar degeneration and retinitis pigmentosa. Although apoB deficiencies result in a very low serum cholesterol level this does not in itself appear to be a major problem - steroid hormone manufacture is normal, except at times of very high demand, and even then women with the diseases can maintain pregnancies (Biemer et al. 1975). However, abnormal membrane fluidity, due to low cholesterol in the plasma membranes does result in misshapen, fragile erythrocytes, a phenomenon termed acanthocytosis (Singer et al. 1952).

1.3.1 Abetalipoproteinaemia

The first of these two diseases, abetalipoproteinaemia, is autosomal recessive in inheritance and patients have no detectable apoB or circulating apoB–containing lipoproteins although their carrier parents have normal levels. In the liver and intestine apoB mRNA levels can be detected, and in some patients they are reported to be increased five-fold over normal

(Lackner et al. 1986) the apoB protein is synthesised but not secreted. In two affected families Talmud et al. (1988) demonstrated that the phenotype did not co-segregate with haplotypes of the apoB gene demonstrating that the cause of abetalipoproteinaemia is not due to mutation in the apoB gene. Sharp et al. (1993) have now demonstrated that the cause, in at least some families, is mutation in the microsomal triglyceride transfer protein (MTP). This gene, which was cloned by the same workers, has no homology with other known proteins and encodes a protein of 894 amino acids (including the signal peptide). Two unrelated patients with abetalipoproteinaemia, both the children of consanguineous marriages, have no MTP activity or protein and have been found to have premature stop codons in their MTP genes, one generating a protein of 78 residues, and the other of 594.

From the study of this disease it has been learned that MTP is necessary for the assembly of apoB containing lipoproteins. MTP adds the lipid to nascent lipoproteins early in assembly within the endoplasmic reticulum, and so rescues the apoB from degradation. MTP also continues to add lipid to the nascent particles up to the point of secretion, and so may also help control the size and density of particle secreted. Studies of these patients have additionally demonstrated that although apoB mRNA appears to be produced constitutively in normal individuals, there must be some feedback control on apoB mRNA, which causes some of these patients to have such high levels.

1.3.2 Hypobetalipoproteinaemia

The second apoB deficiency disease, hypobetalipoproteinaemia has a co-dominant pattern of inheritance. The heterozygote frequency in western populations is estimated to be between 1/500 and 1/1000 (Linton et al. 1993). Heterozygotes have 5-50% of normal levels of apoB and LDL-cholesterol and may be protected from coronary artery disease. The homozygotes have no detectable apoB containing lipoproteins and may be as severely affected as abetalipoproteinaemia

patients. This disease does co-segregate with the apoB gene (Young et al. 1987a, Leppert et al. 1988) and at least 25 different mutations causing hypobetalipoproteinaemia have now been reported (Linton et al. 1993). Almost all of these generate premature stop codons in the apoB gene, although one change in the first nucleotide of intron 5 (Huang et al. 1991) and another in the second nucleotide of intron 24 (Talmud et al 1993) probably interfere with RNA splicing. The phenotype varies depending upon where the truncation occurs. Stop codons 5' of exon 26 generally generate null-alleles: four mutations have been found that would be predicted to produce proteins between 2 and 29 % of full-length, yet these very short apoB's cannot be detected in the plasma of affected heterozygotes. The reason why apoB-25 and apoB-29 are not detected in plasma is unclear, because tissue culture experiments have demonstrated the secretion of proteins as small as apoB-18 from cDNA constructs (Yao et al. 1991, Graham et al. 1991) - there appears to be a mechanism which acts in humans that does not apply in cultured cells. Recently, Talmud et al. (1993) reported a truncated apoB, estimated to be 27.7% of full-length, which can be detected as a lipoprotein in the plasma of both heterozygotes and homozygotes. This species, which is thought to be the product of differential mRNA splicing, is predicted to terminate with a number of hydrophobic residues not normally present in apoB, which may allow sufficient lipid binding to form a lipoprotein particle. Truncations yielding expressed apoB species below 37% of full-length generate lipoproteins with a density similar to HDL, presumably these short apoB proteins cannot bind enough lipid to reduce the density below this. Only 3' stop codons generating proteins larger than apoB-37 are secreted by the liver as particles of VLDL/LDL density range. All heterozygotes have low cholesterol levels from birth but there is no correlation between the size of the apoB protein and serum cholesterol levels (Linton et al. 1993). The effects of LDL-receptor mutations can be ameliorated by hypobetalipoproteinaemia - one individual who is heterozygous for both defects has a normal lipid profile (Emi et al. 1991).

These truncated apoB molecules have been useful in revealing more about apoB

production and clearance, although as yet the picture is still unclear. Hypobetalipoproteinaemia is associated with serum levels of apoB and cholesterol in heterozygotes of only 5-50% of normal, but it has not yet been unequivocally demonstrated whether this is due to underproduction or very rapid removal of the lipoprotein particles - probably both are involved. Ross et al. (1986) found very low levels of hepatic apoB mRNA in a homozygous hypobetalipoproteinaemia patient, indicating that RNA transcription or stability might determine apoB levels in this patient, although this is not the case in normal apoB-100. The shortest apoB species are not detected as lipoprotein in the plasma, perhaps because apoB below about 30% of full-length may be unable to bind sufficient lipid to avoid degradation before secretion. Evidence for a defect in lipoprotein synthesis comes from a patient heterozygous for apoB-46. In this patient an increase in normal apoB-48 chylomicrons was seen in response to a fat-rich meal, but those containing apoB-46 did not rise, although apoB-46 is normally expressed by the intestine of this patient. This indicates that apoB-46 cannot be produced by the intestine in response to dietary fat intake (Young et al. 1990). Fractional catabolic rate (FCR) studies in patients with hypobetalipoproteinaemia have generally indicated decreased LDL production rates but normal clearance (Steinberg et al. 1979, Converse et al. 1989) suggesting that a problem may lie in the metabolism of VLDL to LDL in some patients, however this will be dependant on the initial density range of the lipoprotein containing the truncated apoB, which is in turn dependant on the length of the truncated species.

Several workers have investigated the receptor binding properties of the different truncated apoB species in order to understand the clearance of lipoproteins in hypobetalipoproteinaemia. ApoB-37, which does not have the receptor binding domain, as expected does not bind to the receptor (Young et al. 1987c). However apoB-75, which terminates at residue 3386 and so has only part of the primary receptor binding domain, binds with higher affinity than normal LDL (Krul et al. 1992). ApoB-87 and apoB-89, which contain the primary receptor binding domain but lack the carboxyl-terminal residues, also have higher affinity for the

LDL-receptor (Krul et al. 1989, Gabelli et al. 1989). The results of these studies are further complicated by the fact that a higher proportion of VLDL containing apoBs -75 and -89 are cleared directly, and less progresses to IDL and LDL than is found with apoB-100. There are two possible explanations for this, either the receptor binding domain is exposed in the correct orientation in these VLDLs to be recognised by the LDL-receptor - this would be quite different from apoB-100. Alternatively, these particles may be enriched in apoE and cleared using apoE as the receptor ligand. This later possibility is likely to be the explanation for the clearance of apoB-50 which has no receptor-binding domain. A homozygote for apoB-50 was studied by Hardman et al. (1991), who found that VLDL from this patient were rich in apoE, via which they were cleared rapidly and directly.

In conclusion, the reasons for reduced serum apoB and cholesterol levels in hypobetalipoproteinaemia patients are still not completely clear. There is evidence for both reduced production and enhanced clearance of particles containing truncated apoB species. There are differences depending on the length of the apoB, the shortest species being clearly associated with reduced secretion and lipoprotein production, and on the interaction of the truncated form with the apoB-100 in heterozygotes, as well as interaction with apoE. The reason why heterozygotes, with one normal apoB allele, may have only 5% of normal apoB levels is hard to explain without invoking such mechanisms.

The truncated apoB molecules that cause hypobetalipoproteinaemia have also been of use in understanding the relationship between apoB and apo(a). Coleman et al. (1990) have demonstrated two cysteines (Cys_{3734} and Cys_{4190}), on the surface of the lipoprotein, that are not involved in internal di-sulphide bridge formation within apoB and so have the potential to form a covalent bond with apo(a). Linton et al. (1993) investigated a compound heterozygote with apoB-37/apoB-86 for the presence of Lp(a) - apoB-37 has neither Cys residue and apoB-86 lacks

Cys_{4190} - and found that although this individual did express apo(a), it did not appear to be attached to either truncated apoB species, indicating that they both lacked the attachment site. By a process of elimination, this suggests that Cys_{4190} is the likely site of the di-sulphide bridge linking apoB and apo(a) to form Lp(a) particles.

1.3.3 Familial combined hyperlipidaemia and related syndromes.

Familial combined hyperlipidaemia (FCHL) was first described by Goldstein et al. (1973) as a monogenic disorder in which some family members have hypertriglyceridaemia, others hypercholesterolaemia and some have both. Brunzell et al. (1983) noted that the phenotype could vary within one individual over time. More recently, it has become apparent that FCHL is one of a collection of related syndromes that include hyperapoB (Sniderman et al. 1990), LDL subclass pattern B (Austin et al. 1986, 1988), familial dislipidaemic hypertension (FDH) (Williams et al. 1988) and syndrome X (Reaven et al. 1988). Although there is clearly overlap between these disorders, the exact relationship between them is not clear. FCHL has been reported to affect 0.5-2% of the general population (Goldstein et al. 1973, Nikkila et al. 1973, Grundy et al. 1987) and more than 10% of male survivors of myocardial infarctions under 60 years of age (Goldstein et al. 1973).

HyperapoB, which was originally defined as a high ratio of apoB to cholesterol in LDL (Sniderman et al. 1980) and FCHL have in common an increased number of small, dense LDL particles in the plasma. Every LDL particle has one molecule of apoB but the amount of cholesterol per particle can vary (Teng et al. 1983) and these particles have lower cholesterol to apoB ratio than normal. Grundy et al. (1987) proposed that the characteristic common to all family members with FCHL is overproduction of apoB-100 and increased secretion of apoB containing lipoproteins by the liver, then Sniderman et al. (1990) proposed that FCHL and

hyperapoB are identical syndromes. Small, dense LDL particles are also a defining feature of LDL subclass pattern B (Austin & Krauss 1986). These workers divided LDL into two major sub-classes; large, pattern A with a peak diameter of 26.6nm and small, pattern B with a diameter of 24.8nm. These workers have defined pattern B LDL as atherogenic lipoprotein phenotype (ALP) (Austin et al. 1990b) and have found the frequency for the allele for pattern B to range from 0.19-0.3 (Austin 1993). In families with FCHL affected family members with pattern B LDL have higher serum levels of apoB and triglycerides than affected members with pattern A. Pattern B LDL thus appears to make the phenotype more severe but is not the direct cause of FCHL (Austin et al. 1990a). A second study of FCHL families demonstrated a bi-modal distribution of apoB levels within the group of individuals with pattern B LDL indicating that a second factor might be involved in determining apoB levels in these families (Austin et al. 1992).

The other two syndromes related to FCHL involve hypertension and diabetic factors as well as dislipidaemia. FDH was defined by Williams et al. (1988) as families in which two or more individuals had early onset hypertension in addition to one or more lipid abnormalities; elevated LDL cholesterol, elevated triglyceride levels or low HDL cholesterol. Twin studies demonstrated a genetic basis for this syndrome (Selby et al. 1991) and also showed that FDH twins were obese, glucose intolerant and hyperinsulinaemic and had increased CAD. Hunt et al. (1989) investigated the relationship between FCHL and FDH and found that one third of patients diagnosed as having FDH also fitted the criteria for FCHL diagnosis. Syndrome X was proposed by Reaven (1988) to involve insulin resistance, glucose intolerance, and hyperinsulinaemia as well as hypertension, increased levels of VLDL and decreased HDL levels and so may be related to FDH. It has since been shown that 40% of patients with hypertriglyceridaemia and hyperapoB also have diabetes mellitus and almost 60% have hypertension (Kwitero vich et al. 1993), further linking these syndromes. It is also notable that secondary hyperlipidaemias, for example those associated with diabetes mellitus or chronic renal failure are also associated with over-production

of apoB (Grundy et al. 1987).

In view of the complexity of these syndromes it now seems unlikely that Goldstein et al. (1973) were correct in suggesting that FCHL was a monogenic disorder. Kwitero vich (1993) propose a "two-hit hypothesis" in which mutations in genes generating high levels of apoB interact with mutations in other genes causing hypertriglyceridaemia to produce FCHL and related phenotypes. Candidate genes for FCHL that have been investigated are those for lipoprotein lipase [LPL] (Babirak et al. 1989, Seed et al. 1993), the apoAI/CIII/AIV gene cluster (Wojciechowski et al.1991 Wijsman et al. 1992, Xu et al. 1993), and receptors for acylation stimulating protein and other basic proteins (Cianflone et al. 1990, Kwiterovich et al. 1990). Although the gene for apoB has been investigated with regard to FCHL there is no evidence that it is involved (Rauh et al.1990, Austin et al. 1991, Coresh et al.1992). In addition genes on chromosome 19, including the LDL-receptor and the insulin receptor, have been reported to be linked to LDL subclass pattern B (Nishina et al. 1992).

The reports of FCHL being linked, or not, to various gene loci only begin to address the explanation of the biochemical mechanisms that cause FCHL. If small, dense LDL particles are the unifying feature of the syndromes, these could be due to the rate of secretion of VLDL, the composition at secretion or the metabolism after secretion - the hydrolysis of triglycerides, the exchange of cholesterol ester between cholesterol-rich and triglyceride-rich particles or the rate of removal of LDL.

Babirak et al. (1989) demonstrated a possible link between heterozygous LPL deficiency and FCHL. Williams et al. (1992b) demonstrated that LPL reduces the output of apoB from HepG2 cells, newly secreted apoB encounters LPL in the space of Disse and is promptly taken back up by the hepatocytes in a futile cycle - low LPL levels could reduce the re-uptake and

increase the secretion rate of apoB. Microsomal triglyceride transfer protein (MTP - Section 1.3.1) could also be involved at the stage of apoB secretion. Newly synthesised apoB is rescued from degradation by MTP which adds triglyceride to the nascent lipoprotein, and even after degradation has been prevented, MTP continues to add lipid to the growing particle. Obviously defects in these two roles could lead to a failure in apoB degradation and secretion of particles with incorrect density. In addition to triglyceride transfer by MTP, cholesterol synthesis in the rough endoplasmic reticulum also has a regulating role in the number of apoB molecules that are incorporated into lipoproteins and secreted or are degraded (Cianflone et al. 1990).

Other factors involved in controlling LDL density may act through uptake mechanisms at peripheral locations rather than directly in the liver. Sniderman and co-workers have developed the hypothesis that reduced fatty acid uptake by peripheral cells may result in partially hydrolysed chylomicrons and fatty acids returning to the liver. The raised fatty acid levels, thus generated in the liver, would lead to increased hepatic secretion of apoB (Sniderman et al. 1990). In support of this theory they have demonstrated that after an oral fat load even normotriglyceridaemic hyperapoB individuals clear triglycerides more slowly than normal people (Genest et al. 1986) and that the adipocytes of such individuals synthesise triglycerides more slowly (Teng et al. 1986). These findings also provide evidence that this disorder is multifactorial, individuals may have the underlying predisposition (ie hyperapoB and slow triglyceride clearance) whilst remaining normolipidaemic, indicating that additional genetic or environmental factors are required to cause hyperlipidaemia. Cianflone and co-workers (1989a, 1989b) have discovered a basic protein, called acylation stimulating protein (ASP), Kwitero vich and his colleagues (1990) have additionally discovered two others, which moderate triglyceride synthesis in peripheral adipose tissue and thus affect the rate of triglyceride clearance from plasma (Cianflone et al. 1989a,1989b). There are reports that skin fibroblasts from hyperapoB patients do not respond to ASP as much as fibroblasts from normal individuals (Cianflone et al. 1990,

Kwiterovich et al. 1990) indicating that differences in a putative receptor for these basic proteins may also have a role in FCHL.

Whether the apoAI/CIII/AIV gene cluster is a FCHL locus is still in debate, Wojciechowski et al. (1991) found strong linkage in seven out of sixteen FCHL families (LOD=6.86,θ=0). Additionally, this and other studies have found a higher frequency of a polymorphic XmnI cutting site, in this gene cluster, in patients with FCHL than in controls (Kessling et al. 1985, Shoulders et al. 1986, Hayden et al. 1987). However two recent studies (Wijsman et al. 1992. Xu et al. 1993) have failed to confirm the co-segregation of alleles of the AI/CIII/AIV gene cluster with FCHL. It is biochemically feasible that some of these apoproteins could be involved in FCHL. *In vitro* studies have shown that apoCIII inhibits the hydrolysis of triglyceride from VLDL by LPL (Brown et al. 1972, Wang et al. 1985, Ginsberg et al. 1986) and the clearance of VLDL remnants (Windler et al. 1985). Transgenic mice expressing human apoCIII have hypertriglyceridaemia and elevated cholesterol levels (Ito et al.1990 Aalto-Setälä et al. 1992) and studies with them have indicated that high levels of apoCIII may act as an inhibitor by displacing apoE from lipoprotein particles. ApoAIV may also modulate LPL activity, by facilitating the transfer of apoCII to chylomicrons (Golderg et al. 1990) since apoCII is a co-factor for LPL defects in apoAIV could also have a role in causing FCHL.

Finally, in a study of LDL subclass patten B, Nishina et al. (1992) found strong linkage between ALP and a locus close to or including the LDL-receptor gene (LOD = 4.07, θ=0.04) and moderate linkage with the insulin receptor gene (LOD = 1.78, θ= 0.001). Defects in the LDL-receptor gene are the cause of FH (Brown & Goldstein 1986), however it is possible that certain mutations in the LDL-receptor cause phenotypes different from clinical FH, perhaps by marginally altering the affinity of the receptor for one or both ligands, or by affecting the regulation of expression of the receptor in response to hepatic cholesterol levels. Thus, the LDL-

receptor could modulate LDL density although the linkage study does not exclude the hypothesis that the gene causing ALP is separate from but close to the LDL-receptor on chromosome 19. The gene for the insulin receptor is also in this chromosomal region and, although exhibiting insignificant linkage to pattern B LDL, could also be envisaged to be involved in ALP and the related syndromes involving diabetes. Much work is still required before the molecular bases of this group of related syndromes are understood. Again they illustrate the multifactorial and polygenic nature of hyperlipidaemia and their study, thus far, has revealed more about the processes of lipid metabolism.

1.4 ApoB as a risk factor in coronary artery disease.

The role of apoB in lipid metabolism has made it an obvious candidate gene for the development of coronary artery disease (CAD).

1.4.1 Risk factors for coronary artery disease.

The causes of CAD are multifactorial and polygenic. The factors involved in increased risk of disease include increasing age, type A personality, smoking, high dietary saturated fat intake, high blood pressure, high levels of some blood clotting factors (eg fibrinogen and factor VII), high levels of Lp(a) and high plasma levels of triglycerides and LDL cholesterol. Other factors such as oestrogens and high levels of HDL cholesterol are thought to be protective.

When the coronary arteries of patients who have CAD are examined post-mortem, they are characterised by multiple lesions, called atherosclerotic plaques. The plaques can reach the size where they almost totally occlude the artery, causing angina pectoris. Certain plaques become fragile and rupture, initiating the clotting cascade and causing myocardial infarction if the resultant thrombus occludes the artery. The first stage of plaque formation is the adhesion

of monocytes to the endothelium of the artery, possibly in response to endothelial injury (Ross 1986, for review of atherogenesis). The monocytes then migrate through the endothelium into the arterial wall, where they accumulate and take up cholesterol to become foam cells. At this stage the lesion is called a fatty-streak and is thought to be reversible. If the plaque continues to grow, the smooth muscle cells of the arterial wall proliferate over the cholesterol-filled foam cells and the endothelium may be lost. Eventually the plaque becomes fibrous and calcified, and the cells at the core containing the cholesterol deposit become necrotic. This very complex, multi-step process takes many years to develop, however the cholesterol deposit remains a central feature of atherosclerosis and much research has been concerned with this.

Oxidised LDL is now known to mediate effects at all stages of atherogenesis. Not only is it taken up by the macrophage scavenger receptor during foam cell formation (Section 1.2.6), it has also been found to increase monocyte binding to the endothelium and increase levels of monocyte and smooth muscle cell chemotactic and proliferative factors; interleukin-1β and platelet derived growth factor. In addition it promotes thrombus formation and the toxicity of oxidised LDL increases cell death in the necrotic core of mature atherogenic plaques (Berliner et al. 1993). High levels of Lp(a) also have atherogenic effects mediated both through delivery of cholesterol to the arterial wall and through its adverse effect on the clotting cascade (Section 1.2.5). The affinity of apoB for arterial proteoglycans (Section 1.2.4) may also encourage apoB to stick to the arterial endothelium and initiate atherogenesis.

In addition to the effects of modified lipoproteins (above) high plasma levels of cholesterol, carried as native LDL or VLDL, are a major risk factor in the development of CAD. There is a correlation between increasing plasma cholesterol levels and risk of CAD, however the relationship is not linear. Grundy (1986) used the combined results from three studies; the Framingham heart study, the Pooling project, and the Israeli prospective study to show that above

a plasma cholesterol concentration of 5.17mM risk of CAD begins to rise. The Multiple Risk Factor Intervention Trial (MRFIT) (Stamler et al. 1986), studying 350 000 men, did not find this threshold effect, but otherwise found a similar relationship; the relative risk of death from CAD is considered to be 1 at a cholesterol concentration of 5.17mM, it doubles at a concentration of 6.47mM and quadruples at 7.76mM. Other risk factors such as smoking, hypertension, obesity and diabetes interact with plasma cholesterol level in an individual, in an additive or possibly multiplicative manner, to reduce the age of onset of CAD (Grundy 1986). Patients with heterozygous defects in the LDL-receptor (familial hypercholesterolaemia - FH) have only half the normal numbers of LDL-receptors, resulting in slow clearance of LDL-cholesterol and consequently have levels of serum cholesterol above the 95th percentile of the normal distribution. Even in the absence of other risk factors, such patients usually develop CAD and begin to suffer myocardial infarctions from their fourth and fifth decades. Patients who are homozygous for LDL-receptor defects have six to ten times the normal cholesterol levels and usually die in adolescence or early-adulthood (Brown & Goldstein 1986).

Although environmental factors such as diet clearly have a role in determining plasma cholesterol levels, there are also strong genetic factors involved. Much work over recent years has been concerned with elucidating which genes have a role in determining serum cholesterol levels and related risk of CAD. Since there is a positive correlation between plasma apoB levels and LDL-cholesterol levels, the gene for apoB is an obvious candidate for being one of the genetic factors which controls plasma cholesterol levels. Some studies have shown apoB levels to be a better discriminator than lipid levels for risk and degree of CAD (Avagaro et al. 1979, Whayne et al. 1981).

1.4.2 The apoB gene is a candidate gene for determining serum cholesterol levels.

Path analysis and twin studies have shown that serum LDL-cholesterol and apoB levels

have a high heritability of between 0.5 and 0.6 (Hamsten at al. 1986a, Berg et al. 1987a). Family studies using complex segregation analysis have found evidence that plasma levels of apo B are determined by a major gene, with a contribution from environmental factors and genes of small or intermediate effect, which account for 35-55% of the variation in apoB levels (Hasstedt et al. 1987, Pairitz et al. 1988, Coresh et al. 1992). The question remains as to what is the major gene that determines apoB levels. Although several studies have found evidence for apoE genotype having an effect on apoB levels, (Section 1.2.7) Pairitz et al. (1988) conclude that apoE is not the gene with the major effect. Coresh et al. (1992), in a linkage study using 23 families, find no evidence for the apoB gene being the major gene either. However, this last study specifically selected families with large differences in apoB levels, which perhaps resemble FCHL or even FH pedigrees. Such pedigrees would show evidence for a major gene having an effect on apoB levels, but it is known that in neither case is this the apoB gene (Section 1.3.3, Brown & Goldstein 1986). It is possible that there is heterogeneity, with different genes have the major effect in different families. Certainly there is evidence of rare mutations in genes that have large effects on apoB levels, abetalipoproteinaemia (Section 1.3.1), hypobetalipoproteinaemia (section 1.3.2) FCHL (Section 1.3.3) and Familial Defective ApoB-100 (FDB) (Section 1.4.4) are all examples of these. The apoB gene is conclusively implicated in two of these defects: FDB and hypobetalipoproteinaemia. However there is also evidence of common polymorphisms in genes that cause small rises in apoB levels, and predispose to disease only when they interact with other factors. Polymorphism in the apoE gene has this effect (Sections 1.2.7 & 1.2.8) and there is evidence that polymorphism in the apoB gene has a similar effect (Section 1.5.3).

1.4.3 Familial defective apoB-100 (FDB).

Definitive proof of mutation in the gene for apoB causing raised apoB levels and CAD came with the discovery of FDB. A patient with raised cholesterol levels and who had LDL with only 30% of normal binding to the LDL-receptor was identified (Vega and Grundy, 1986,

Innerarity et al. 1987). This patient was found to have a substitution of glutamic acid for arginine at residue 3500 (Soria et al.1989). Lund-Katz et al. (1989) demonstrated that this substitution affects the local conformation of apoB, changing the microenvironments of six normally "active" lysines (pK8.9) to the inactive form (pK10.5) (Section 1.2.3). This charge and conformation change alters the receptor binding domain and virtually abolishes the ability of mutant LDL to bind to a normal receptor - it has only 3-5% of normal binding activity (Innerarity et al. 1988). Many individuals have now been discovered with the same mutation. It has an autosomal dominant mode of inheritance with a heterozygote frequency estimated to be 1/600 in the general populations of North America and Europe, rising to 3-6% among patients with Type IIa hyperlipidaemia or a clinical diagnosis of familial hypercholesterolaemia (FH) (Tybjaerg-Hansen et al. 1992). Most, but not all, of the individuals identified with this mutation have moderate to severe hypercholesterolaemia (Tybjaerg-Hansen et al. 1992). Affected individuals have raised serum LDL- and total cholesterol levels, but VLDL-, HDL-cholesterol and triglyceride levels are essentially normal. Many carriers also have abnormal tissue-cholesterol deposits in the skin, tendons and corneae, typical of FH (Tybjaerg-Hansen et al. 1990, Myant et al. 1991), indicating that mutations in both the receptor and the ligand can generate identical clinical phenotypes.

The vast majority of individuals identified as carrying the mutation encode the glutamine residue on the same haplotype of the apoB gene, indicating that these individuals have a common ancestor (Ludwig and McCarthy 1990, Rauh et al. 1992). Two patients, one in Germany and another from China, have now been identified with the Gln_{3500} encoded on different haplotypes (Rauh et al. 1992, Bersot et al.1993) (Table 1.4.3). These haplotypes do not appear to have arisen by recombination with the original one, but are more likely to represent independent new mutations at the same site. The base substitution at codon 3500 which generates the mutation is CGG to CAG - it involves a CpG dinucleotide and is thus a mutational hot-spot. The existence of individuals with independently occurring mutations on different haplotypes but with the same

phenotype provide strong evidence that this G to A nucleotide substitution, and not some other base-change in linkage disequilibrium with it, is the cause of this form of defective apoB. However, only *in vitro* expression studies will prove this conclusively.

Table 1.4.3 The three different haplotypes on which the Gln_{3500} substitution is encoded.

*	5' $(TG)_n$	Sig. Pep.	71 ApaLI	HincII	PvuII	591 AluI	2488 XbaI	3611 MspI	4154 EcoRI	3' HVR
E1	14	+	+	–	–	–	–	+	–	48
E2	14						+	–	+	30
C	14	+	+	–	–	–	–	+	+	30

E1 = common European haplotype (Ludwig and McCarthy,1990),
E2 = 2nd European haplotype (Rauh et al. 1992),
C = Chinese haplotype (Bersot et al. 1993)
* The polymorphisms used to define the haplotypes are described in section 1.5.2. The residue number affected by the polymorphism is given where appropriate.

A patient homozygous for Gln_{3500} has been reported (Marz et al. 1992). Unlike patients who are homozygous for LDL-receptor mutations, this patient was not more severely affected than the heterozygotes. It is probable that the normal apoE in this patient is able to act as a ligand for clearing lipoprotein particles via the LDL-receptor. In such patients with no apoB functional as ligand for the LDL-receptor, the receptor will not be down-regulated by the clearance of apoB-LDL and apoE will not have a competitor for binding to receptors. This form of defective apoB is thus having a truly dominant effect.

1.5 Common genetic variants of apoB.

1.5.1 Antigen group protein polymorphisms.

Protein polymorphisms of apo B were first detected using antisera from multiply-

transfused patients (Allison & Blumberg 1961). The Antigen Group [Ag] system of variants that these workers reported are caused by different antigens on LDL. There are ten epitopes which were found by Bütler and Brunner (1974) to be arranged in five tightly linked loci, each with two alleles. The epitope pairs are called: Ag(a1/d), Ag(c/g), Ag(h/i), Ag(t/z) and Ag(x/y) (reviewed in Breguet et al. 1990) [**Note**: For simplicity within this thesis the Ag(a1/d) pair is referred to as (a/d)]. Since the cloning of the human apo B gene, several nucleotide substitutions have been reported to be candidates for the molecular bases of the Ag epitopes (Ma et al. 1987, 1989, Dunning et al. 1988, Wang et al. 1988, Young et al. 1989, Xu et al. 1989, Wu et al. 1991, Reviewed in Section 1.5.2) and this confirms that the antigens are generated by protein polymorphisms of apoB expressed on LDL particles. Some monoclonal antibodies raised to apoB recognise epitopes of the Ag system - both MB19 (Tikkanen et al. 1986) and BIP-45 (Duriez et al.1987) recognise the Ag(c) epitope of the Ag(c/g) pair. Schlapfer et al. (1987) have additionally generated a monoclonal that detects the Ag(d) epitope of the Ag(a/d) pair.

1.5.2 DNA polymorphisms.

Since the cloning of the apoB gene at least 20 polymorphisms have been reported within the gene and its flanking regions. These are shown in Table 1.5.2. Initially polymorphisms were detected as Restriction Fragment Length Polymorphisms (RFLPs), these can be single base changes that create or destroy recognition sites for restriction enzymes, or insertions, deletions and repeated elements that alter the distance between restriction sites. More recently, with the advent of techniques for detecting single base changes in genes, polymorphisms that do not alter restriction enzyme sites have also been found. In addition, comparison of the eight published cDNA sequences of the apoB gene have revealed more than 70 potential sequence variants (Ludwig et al. 1987, Yang et al. 1989). However these probably include many sequencing errors and rare variants, as well as true polymorphisms.

Table 1.5.2 Polymorphisms in the apoB gene.

Position	Base(s) affected	Restriction Enzyme	Amino acid change
5′ Flanking HVR [1]	$(TG)_n$ 3256bp 5′ to start	-	-
5′ Flanking [2]	≈ -4kb 5′ to start	AvaII	-
Promoter[3]	C to T at -265bp	MspI	-
Exon 1[4]	9bp deletion		LeuAlaLeu deletion from signal peptide
Intron 3[2]	G to T 92bp 3′ to ex3	-	-
Exon 4[5]	C_{293} to T	ApaLI	Thr_{71} to Ile
Intron 4[6,2]	A to C 171bp 3′ to ex4	HincII	-
Intron 4[6,2]	A to G 523bp 5′ to ex5	PvuII	-
Exon 14[7]	T_{1981} to C	AluI	Ala_{691} to Val
Intron 20[2]	G to A 146bp 5′ to ex21	BalI	-
Intron 20[8]	$(TTTA)_n$	-	-
Exon 26[9]	C_{7673} to T	XbaI	-
Exon 26[10]	C_{8344} to T	-	Pro_{2712} to Leu
Exon 26[11]	C_{11040} to G	MspI	Arg_{3611} to Gln
Exon 29[12]	G_{12669} to A	EcoRI	Glu_{4164} to Lys
Exon 29[Chapter6]	G_{12937} to C	-	Arg_{4243} to Thr
Exon 29[13,14]	A_{13141} to G	-	Asn_{4311} to Ser
3′ Flanking VNTR[15,16]	11-16bp AT rich tandem repeat 181bp 3′ to ex29	MspI, BamHI, EcoRV	-

References: 1- Ludwig & McCarthy 1990, 2 - Huang et al. 1990a, 3 - Jones et al. 1989, 4 - Boerwinkle & Chan 1989, 5 - Young & Hubl 1989, 6 - Darnfors et al. 1986, 7 - Wang et al. 1988, 8 - Zuliani & Hobbs, 1990, 9 - Talmud et al. 1985, 10 - Huang et al. 1990b, 11 - Huang et al. 1988, 12 - Shoulders et al. 1985, 13 - Navajas et al. 1991, 14 - Dunning et al. 1992, 15 - Berg et al. 1986, 16 - Boerwinkle et al. 1989.

Association studies between these DNA polymorphisms and the Ag protein polymorphisms have identified amino acid substitutions as the candidate molecular bases for all the Ag epitopes. These are shown in Figure 1.5.3. Although some of the authors of these reports have claimed that these substitutions are responsible for the Ag polymorphisms, association studies do not warrant such claims, as demonstration of correlation does not prove causation. Only peptide expression studies and the use of antibodies will demonstrate unequivocally the amino acid substitutions that create the Ag epitopes. This is discussed further in Chapters 4 and 7.

In addition to the complete allelic association seen between the DNA and protein polymorphisms shown in Figure 1.5.3, strong linkage disequilibrium and allelic association have been reported between many of the polymorphic sites in the apoB gene. Berg et al. (1986) reported linkage between the Ag system and the apoB gene (LOD = 4.5, θ = 0) and noted the allelic association of the XbaI X- allele with the Ag(x) epitope. There have also been reports of linkage disequilibrium between the XbaI RFLP and the following sites: Ag(c/g), Ag(x/y), Ag(t/z) and the signal peptide insertion/deletion (Ma et al. 1987, Dunning et al. 1988, Berg et al. 1986, Myant et al. 1989, Xu et al. 1990, Renges et al. 1992), the PvuII RFLP with Ag(a/d) (Dunning et al. 1988), and Ag(t/z) with the 3'VNTR polymorphism (Hegele et al. 1986, Renges et al. 1992). These were indications that there is strong linkage disequilibrium and allelic association between polymorphic sites across the entire apoB gene. This is evidence of an evolutionary relationship between these alleles, which is not necessarily based on physical distance between the polymorphic sites. This topic will be extended and further discussed in Chapters 5 and 6.

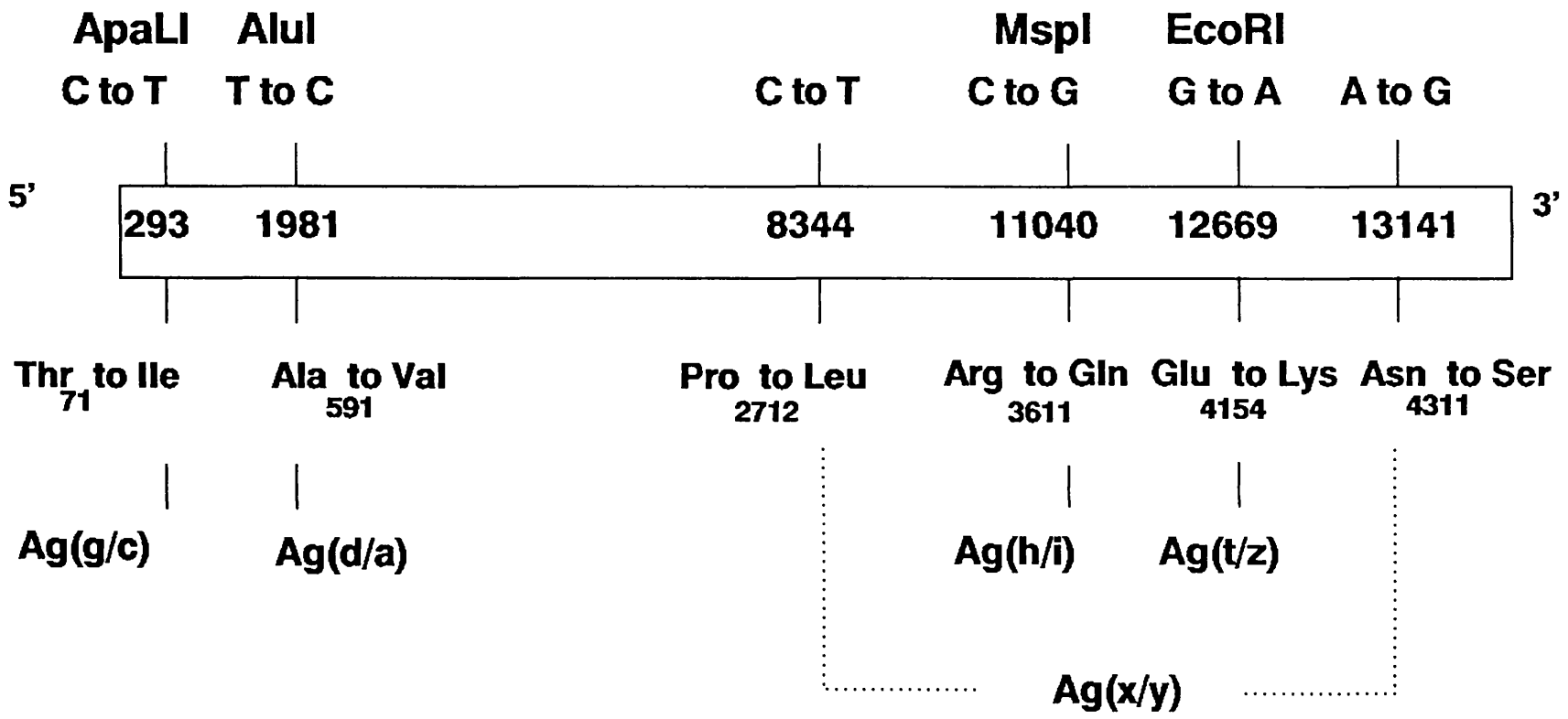


Figure 1.5.3 The candidate molecular bases of Ag polymorphisms

Changes affecting the DNA are shown above the box (where nucleotide numbers are shown) and changes affecting protein are shown below . RFLP sites are marked along the top and Ag polymorphisms along the bottom.

Table 1.5.4 Findings of published association studies investigating the X+ allele of the XbaI RFLP and serum lipid levels in healthy and random population samples.

Reference	Sample size	Age range(years)	Total Chol.	LDL Chol.	apoB	Trigs.
Berg 1986	56 MF	20 - 29	↑** 12%	-	↑** 19%	↑
Hegele et al. 1986	84 MF	58 ±10	↑	↑	↑	↑
Law et al. 1986	83 M	60 ± 7	↑** 7%	-	↑	↑** 39%
Talmud et al. 1987	62 MF	20 - 65	↑** 19%	↑* 30%	-	↑** 47%
Aalto-Setälä et al. 1988	176 MF	20 - 66	↑** 11%	↑* 12%	↑	↑
Arburatani et al. 1988	54 MF	19 - 62	↓*	-	↓	↓
Dunning et al. 1988	51 MF	43	↓*	-	↔	↔
Jenner et al. 1988	22 MF	25 - 40	↑	-	-	↑**
Darnfors et al. 1989	187 MF	27 - 65	↑	↑	↔	↓
Myant et al. 1989	73 M	25 - 60	↑	-	-	-
Myklebost et al. 1990	105 M	43	↑	-	-	↑
Paulweber et al. 1990	118 M	50 ± 5	↑	↑	-	↓
Renges et al. 1991	107 M	40 - 69	↑* 10%	-	-	↑
Deeb et al. 1991	162 MF	⟨ 65	↑** 7%	↑** 10%	↑** 15%	-
Peacock et al. 1992	83 M	23 - 44	↑	↑** 23%	↑	↑
Saha et al. 1992a	196 MF	25 - 60	↓	↓	↓	↑
Saha et al. 1992b	154 M	44 ± 1	↓	↓	↑	↑
Vilella et al. 1992	228 M	19 - 65	↑** 9%	↑** 11%	↑** 17%	-
Friedlander et al. 1993a	525 MF	29 ± 6	↑** 5%	↑** 8%	↑** 8%	↑
Hansen et al. 1993	464 M	45 ± 0	↑** 7%	-	↑** 8%	-

KEY: M = males, F = females, ↑ = raised, ↓ = lowered, ↔ = no effect, - = not determined, * = p ⟨ 0.1, ** = significant p ⟨ 0.05, Percentage values are the difference in mean level between X+X+ homozygotes and X-X- homozygotes.

1.5.3 Associations of the polymorphisms with lipid levels and CAD.

The first finding of association between alleles of the apoB gene and differences in serum lipid levels came from Berg et al. (1976). Using the combined data from ten different populations, these workers reported that the Ag(x-) allele [equivalent to Ag(y)] is associated with significantly higher mean serum cholesterol and triglyceride levels, although the effect is small. The population samples comprised male and female Scandinavians and Americans of European and African descent, aged from 16 years to elderly, with and without CAD. These samples were analysed separately and combined (totalling 1087 individuals). No significant associations were seen among the samples of young individuals, but when the older samples were pooled the Ag(y) allele was associated an average of 4% higher serum cholesterol levels and 13% higher triglyceride levels. Although the researchers reporting this association described the Ag(y) allele to be associated with raised lipid levels, the converse: that the Ag(x) allele is associated with reduced lipid levels, is also seen in this study. This will be discussed further in Chapter 4.

Since the discovery of the XbaI RFLP numerous studies have documented the association of the presence of the XbaI cutting site [the X+ allele] with raised levels of serum total- and LDL-cholesterol, apoB and triglycerides. The findings of 20 published studies on random population samples or healthy individuals are tabulated in Table 1.5.4. Sixteen of the studies found an association of the X+ allele with raised total cholesterol levels (eight of these are significant), ten also found an association with raised apoB levels (five significant), and twelve with raised triglyceride levels (three significant). In addition to the associations seen with the XbaI RFLP, the signal peptide insertion/deletion polymorphism has been shown to be associated with differences in serum triglyceride levels (Xu et al. 1990, Renges et al. 1991), and with response of serum lipids to dietary change (Xu et al. 1990) - individuals homozygous for the insertion allele have the highest triglyceride levels on a normal diet and the greatest response to

a lipid lowering diet. Other studies have also noted the allele encoding the Ag(g) epitope to be associated with lower cholesterol and apoB levels (Tikkanen et al. 1988), and the V- allele of the PvuII RFLP to be associated with raised cholesterol levels (Myant et al. 1989).

In view of the allelic association reported between the Ag(x) and XbaI X- alleles (Ma et al. 1987, Dunning et al. 1988), the studies in Table 1.5.3 and that of Berg et al. (1976) may be investigating the same effect - this will be discussed further in Chapters 4 and 5. Most of the published studies report a trend for the X+ allele to be associated with raised levels, but the effect is small. Estimates for the amount of phenotypic variance in serum cholesterol and apoB levels explained by the XbaI RFLP range from 1 - 21% (Talmud et al. 1987, Friedlander et al. 1993a, Hansen et al. 1993). When the X+X+ homozygotes and the X-X- homozygotes are compared, there is a 7 - 19% difference between total cholesterol and apoB levels in those studies that achieve statistical significance (Table 1.5.3). This difference is smaller in the non-significant studies, however these smaller differences may be valid results - it is necessary to have a very large sample in order to demonstrate a small significant difference. Publication bias against negative results may have resulted in some non-significant results remaining unpublished. None-the-less, the consensus is that the X+ allele is associated with a small rise in lipid levels, but the effect may be masked by the effects of other genetic and environmental factors which also modulate lipid levels. Such factors include age, sex, body-mass-index and secondary causes of hyperlipidaemia. As a consequence of these, it is often more difficult to demonstrate the association in groups of CAD patients than in normal individuals (Berg et al. 1976, Hegele et al. 1986, Monsalve et al. 1988, Myant et al. 1989, Genest et al. 1990), probably because the patients already have more confounding factors which result in a pronounced rise in lipid levels. Evidence in support of this comes from Friedlander et al. (1993b) with a dietary intervention study which could detect an XbaI genotype difference when the study individuals were on a low-saturated-fat diet, and had lower serum lipid levels, but not on the high-saturated-fat diet when

lipid levels were raised. Furthermore, if the causes of CAD in a sample are carefully controlled, the effect of apoB genotype on lipid levels can be detected: Aalto-Setälä and co-workers (1989) studied 120 patients with heterozygous FH and found that the individuals homozygous for the X+ allele had 14% higher total cholesterol and 21% higher LDL-cholesterol levels than the individuals homozygous for the X- allele. These differences are comparable with the differences in normolipidaemic samples even though the FH patients had mean serum cholesterol levels at least twice as high as the normals. Thus it can be concluded that the effects of variance in the gene for the LDL-receptor and apoB are additive.

Other studies have taken the alternative approach of comparing apoB allele frequencies between matched groups of patients and controls. Several studies have found a higher frequency of the XbaI X- and EcoRI R- alleles amongst patients with CAD and peripheral artery disease than in controls (Hegele et al. 1986, Monsalve et al. 1988, Myant et al. 1989, Genest et al. 1990, Paulweber et al. 1990, Renges et al. 1991). However, as stated above, no association was found with serum lipid or lipoprotein levels in most of these studies and in any case the X- allele is associated with lower lipid levels in the studies where an effect is seen. Thus it would appear that the effects of these alleles on CAD risk are mediated through a mechanism other than raised cholesterol levels. One possibility could be that CAD patients carrying the lipid raising X+ allele are under-represented because they have already died, but this would seem unlikely since the lipid raising effect of the X+ allele is so small. The Gln_{3500} allele which causes a very much larger rise in serum cholesterol levels is found more frequently amongst CAD patients than among the normal population (Tybjaerg-Hansen et al. 1990), and the available evidence does not support the idea that the effect of the X+ allele on patient mortality is more severe than the effects of the Gln_{3500} allele.

1.5.4 Possible mechanisms for the association of the X+ allele with raised lipid levels.

Raised serum cholesterol and apoB levels could be the result of over-production or under-removal of LDL particles. Four studies have addressed the question of which of these mechanisms is acting to generate the differences seen associated with the apoB XbaI RFLP. Houlston et al, (1988) looked at the association of the XbaI RFLP with fractional catabolic rate [FCR] of LDL in 22 individuals using autologous LDL labelled with ^{125}I with a urine/plasma ratio method of detecting FCR. Individuals homozygous for the X+ allele had 13-16% higher serum cholesterol and LDL-cholesterol levels and 11% lower LDL FCR than individuals homozygous for the X- allele. In the X+ individuals, the lower FCR was due to both higher production rate of LDL and lower degradation rate than in the X- individuals. Demant et al. (1988) obtained similar results in a study of 19 subjects with moderate hypercholesterolaemia using a slightly different method. Here autologous LDL was labelled with both ^{125}I and ^{131}I, the ^{131}I fraction was then modified at the arginine residues to destroy its ability to interact with the LDL-receptor - in this way both LDL-receptor-mediated and receptor-independent catabolism could be compared. This study found the individuals homozygous for the X+ allele to have an 18% lower LDL FCR than the X-homozygotes. This difference was attributable to LDL-receptor-mediated clearance, which was 39% lower in the X+ homozygotes, whilst there was no genotype difference on clarence by receptor-independent pathways. No association of lipid levels with apoB genotype were seen in this study, but this is not surprising since all the study individuals were moderately hypercholesterolaemic.

The results of these turnover studies are further supported by an *in vitro* study investigating LDL degradation by dermal fibroblasts (Series et al. 1989). In this study LDL was isolated from 33 individuals homozygous for either the X+ or X- allele and labelled with either ^{125}I or ^{131}I. The relative degradation rates of pairs of X+ and X- LDL by fibroblasts were then

compared. No differences were found in LDL from young, normolipidaemic individuals, but when LDL form older or moderately hypercholesterolaemic individuals was investigated, X+ LDL was catabolised 15 - 20% slower than X- LDL. However, another study by Vilella and colleagues (1992) made contrary findings in *in vitro* assays of the LDL and LDL-receptors from 30 individuals (10 of each XbaI genotype). No differences in LDL-receptor activity were found between the genotype classes, but when LDL from the individuals was bound to receptors on a fibroblast cell-line X+ LDL had significantly higher maximum binding (B_{max}) than X- LDL. However the individuals investigated in this study were young and normolipidaemic and did not exhibit the XbaI genotype related lipid and lipoprotein differences seen in the larger sample from which they were drawn. Overall, these results indicate that the raised lipid levels seen associated with the X+ allele are due to slower LDL-receptor mediated catabolism of X+ LDL. Thus the raised levels are due to under-removal rather than over-production of apoB-LDL.

Dietary intervention studies (Tikkanen 1990, Xu et al. 1990,Friedlander et al. 1993b) have begun to investigate the interaction of apoB XbaI genotype with environmental factors (diet). In a cross-over study using low and high saturated fatty acid diets (Friedlander 1993b) the serum apoB, LDL and total cholesterol levels of the X- homozygotes changed more in response to diet than the levels in the X+ individuals. Thus, in response to the high saturated fat diet, serum lipid levels rose in both groups but the magnitude of the rise in the X- homozygotes was more than twice that in subjects with an X+ allele. This study was carried out in 20 young, normolipidaemic, offspring of MI patients, whose normal diet is low in saturated fatty acids and may provide clues as to why the XbaI X- allele is sometimes seen to be more frequent among CAD patient groups than among controls. Although individuals with the X- allele normally have lower cholesterol levels, the interaction between this allele and a high-saturated-fat diet may have a more serious effect on serum lipid levels than is seen with the X+ allele. However these results are not supported by those of another dietary intervention study on 103 individuals from

Finland whose normal diet is high in saturated fat. In this sample the greatest response to a lipid-lowering diet was seen in individuals carrying the apoB XbaI X+ allele and/or the insertion allele of the signal peptide polymorphism (Tikkanen et al. 1990, Xu et al. 1990). The differences in the results of these two studies may be due to their investigating different ethnic groups with different normal diets, however both indicate a relationship between apoB variants and response to dietary fat intake.

The XbaI RFLP cannot be the direct cause of any of the effects seen, since the base change that creates the variant XbaI site is in the third base [wobble position] of codon 2488 and so does not alter the threonine residue at this position (Carlsson et al 1986). The alleles of this RFLP must be in linkage disequilibrium with another base change, or changes, that do have a functional effect leading to the reported differences in lipid parameters. The base changes that do have a direct effect could generate amino acid substitutions affecting the structure of the apoB and thereby its function. Alternatively base changes in a control region of the gene, such as the promoter or enhancer regions, could affect the rate of mRNA transcription. This latter possibility is unlikely in this case, not only because of the data suggesting that the problem lies with clearance rather than production, but also because apoB mRNA transcription does not appear to be a controlling factor of apoB levels in normal individuals. However it is possible to postulate mutations that would affect rate of processing or degradation of nascent apoB - for example the signal peptide insertion/deletion polymorphism may indeed alter the processing rate in a model system (Sturley et al. 1993).

1.5.5 The intentions of the studies described in this thesis.

In the light of the studies suggesting that the apoB XbaI X+ allele is associated with slow receptor-mediated LDL clearance, it was decided to test the hypothesis that the XbaI RFLP was in linkage disequilibrium with another common variant of the apoB gene that altered the amino

acid sequence of the apoB primary receptor binding domain. Such a mutation could affect the charge or tertiary structure of the domain and change the affinity of the apoB for the LDL-receptor, so changing the clearance rate. This would be analogous to the effects of the different isoforms of apoE. By contrast, the Arg_{3500} to Gln substitution in apoB, which clearly alters the charge and conformation of the receptor-binding domain and destroys affinity for the LDL-receptor, is a rare mutation in the general population with a much more serious effect on cholesterol levels than the alleles of the XbaI RFLP. Chapters 3, 4 and 6 of this thesis deal with the search for genetic variations that have the potential to alter the primary or secondary receptor binding domains of apoB. Although this has been the emphasis in this thesis, it must be remembered that amino acid substitutions in apoB could also have other effects on hyperlipidaemia and atherosclerosis, for instance by changing the affinity of apoB for arterial proteoglycans, or altering the rate of oxidation of LDL. Chapter 5 deals with the relationships of polymorphisms of apoB in haplotypes. Having found variants and haplotypes with an effect on the function of apoB, expression studies will be necessary to test the mechanism of action of the variant forms. Chapter 7 deals with the first stages of the *in vitro* expression of peptides of apoB, with the intention of eventually devising functional assays for different forms of apoB.

CHAPTER 2

SUBJECTS, METHODS AND MATERIALS

SUBJECTS

2.1.1 Individuals used in the search for new variants.

DNA from ten hypercholesterolaemic patients was examined in the study described in Chapter 3. Six of these (patients 1 to 5 and patient 9) were drawn from the Glasgow sample (Section 2.2.1.) and two (patients 6 and 7) from the Sheffield sample (Section 2.2.2). Patient 8 from Stockholm, Sweden is a 51 year old man who suffered myocardial infarctions at the age of 43 and 47 and an iliaco-femoral artery thrombosis at the age of 49. FH was excluded in this patient on the grounds of normal LDL-receptor number and function on the patient's cultured lymphocytes. Patient 9 was used as the negative control in Chapter 3. He is homozygous for the XbaI X- allele and has a high FCR of LDL. Patient 10 is heterozygous for the Arg_{3500} to Gln substitution and was used as a positive control for a single base change in the gene region being investigated. He was discovered and is described in more detail in the study of Tybjaerg-Hansen et al. (1990). Relevant characteristics of these patients are given in the top part of Table 2.1.1. With the exception of patient 10, none of the patients in Table 2.1.1 have FDB-100 (Section 1.4.3).

DNA from 12 individuals was investigated by Dr. H Renges in a study described in Chapter 6. LDL from these individuals had been tested in the U937 assay (Frostegård et al. 1990). Seven of these patients had significantly reduced binding of their LDL to a normal receptor, six drawn from the Sheffield study (Section 2.2.2), together with patient 8 (above). Five had increased binding as measured in the U937 assay, they had also been drawn from the Sheffield sample (Section 2.2.2). Characteristics of the individuals investigated in this study are given in Table 2.1.1 (individuals 7,8,9 and 11 to 19).

Table 2.1.1 Characteristics of the individuals investigated for new variants of apoB in Chapter 3(top) and Chapter 6 (bottom of table) .

Individual	Total Chol. (mmol/l)	Triglycerides (mmol/l)	LDL Chol. (mmol/l)	ApoB (mg/100ml)	FCR[d] (pool/day)	U937 cell growth (%)	XbaI Genotype
1[a]	8.33 ±0.22	2.23 ±0.40	5.46 ±0.07	253	0.218[e]	-	X+X+
2[a]	8.56 ±0.47	2.76 ±0.84	6.20 ±0.60	175	0.197[e]	-	X+X+
3[a]	6.36 ±0.19	0.82 ±0.05	3.76 ±0.22	146	0.182[e]	-	X+X+
4[a]	8.00 ±0.38	3.11 ±0.33	5.62 ±0.39	176	0.265[e]	-	X+X+
5[a]	7.03 ±0.20	1.68 ±0.19	4.88 ±0.24	160	0.217[e]	-	X+X+
6[g]	9.32	3.91	5.79	-	0.258	43/17/35[f]	X+X+
7[g]	7.61	3.27	5.00	-	0.302	42/68/74[f]	X+X+
8[g]	7.09	1.01	4.86	-	-	49/43[f]	X+X-
9[ab]	7.08 ±0.15	1.43 ±0.26	5.28 ±0.36	179	0.297[e]	-	X-X-
10[c]	8.0	0.80	6.2	-	-	22/25	X+X-
11	7.64	7.18	4.28	-	0.223	141	X+X+
12	6.37	1.34	4.65	-	0.258	131	X-X-
13	5.94	1.23	4.18	-	0.282	127	X+X+
14	8.49	1.44	6.67	-	0.196	119	X-X-
15	4.87	2.01	2.22	-	0.390	114	X-X-
16	4.64	1.37	2.57	-	0.339	56	X+X-
17	8.12	6.35	4.42	-	0.452	61	X-X-
18	5.02	2.23	3.02	-	0.423	72	X+X-
19	8.41	3.48	5.70	-	0.272	72	X+X+

Footnotes to Table 2.1.1

a - Lipid levels are mean (±SD) of three determinations.

b - Patient used as negative control and probe in chemical mismatch analysis.

c - Carrier of Gln_{3500} mutation, used as positive control in the chemical mismatch analysis in Chapter 3.

d - Mean FCR for entire Sheffield sample = 0.325 ± 0.07 pools /day.

e - Patients from Glasgow sample (Demant et al. 1988), FCR not directly comparable with other samples. Mean for this group = 0.350 ± 0.06 pools/day.

f - Results of repeat determinations on samples taken one month apart are separated by /.

g - Samples used in the studies described in Chapter 3 and Chapter 6.

2.2 Population samples screened for apoB variants.

2.2.1 Glasgow

A group of 17 unrelated hyperlipidaemic individuals on whom fractional catabolic rate of LDL studies have been determined by *in vivo* turnover studies of autologous LDL (Shepherd et al., 1979). Familial hypercholesterolaemia was excluded on the basis of standard criteria, and the patients have been described in Demant et al. (1988) (Section 1.5.4).

2.2.2 Sheffield

A group of 22 unrelated men on whom fractional catabolic rate of LDL had also been performed (Turner et al. 1984, International Collaborative Study Group, 1986). Familial hypercholesterolaemia was excluded on the basis of standard criteria, these individuals have been described in Houlston et al. (1988) (Section 1.5.4).

2.2.3 North Karelia , Finland

A group of 89 unrelated men and women, with no history of hyperlipidaemia, from a rural community in North Karelia, Finland. This population has been very stable for several hundred years with no influence of Swedish speaking individuals. They have been described in detail elsewhere (Ehnholm et al. 1982, 1984, Kuusi et al. 1985). Ag phenotyping, and apoB RFLP genotyping have also been previously described (Xu et al. 1989, Tikkanen et al. 1989).

2.2.4 Finnish

A second group of 24 Ag phenotyped, unrelated Finnish individuals was also examined. Phenotyping and apoB genotyping has been described by Dunning et al. (1988).

2.2.5 Swedish controls

A group of 90 randomly selected, healthy, male residents of Stockholm county, age-matched with the

patient group described below. These individuals were free of symptoms and clinical signs of CAD. Data on lipid, lipoprotein, apoprotein, haemostatic and metabolic variables have been presented elsewhere (Hamsten et al. 1986b,1987). ApoB genotyping has been described in Peacock et al. (1992). One individual with a serum triglyceride level > 3.0mM was excluded from the lipid and lipoprotein analyses. Mean serum levels of certain lipid traits and apoB together with the amount of phenotypic variance explained by age and BMI are given in Table 2.2.5.

Table 2.2.5 Mean serum lipid and lipoprotein traits and phenotypic variance explained by age and BMI in the Swedish control group.

	Mean (± SD) serum level	Phenotypic variance explained by differences in age and BMI
Total Cholesterol mmol/l	6.11 ± 1.18	3.0%
Total Triglyceride mmol/l	1.51 ± 1.36	13%
LDL Cholesterol mmol/l	3.99 ± 0.93	3.7%
HDL Cholesterol mmol/l	1.42 ± 0.34	13%
ApoB mg/100ml	108.27 ± 18.9	4.5%

2.2.6 Swedish patients

The patient group consisted of 143 male and female individuals from Stockholm county, Sweden, who had suffered a myocardial infarction below the age of 45 years. Recruitment details, data on lipid, lipoprotein, apoprotein, and metabolic variables and descriptions of DNA preparation for these individuals have been presented elsewhere (Hamsten et al. 1986b,1987). ApoB genotyping has been described in Peacock et al. (1991). Individuals with serum cholesterol levels > 9.5mM, and/or triglyceride levels > 3.0mM, or with a diagnosis of familial hypercholesterolaemia, diabetes or porphyria were excluded from

the lipid and lipoprotein analyses. Mean serum levels of certain lipid traits and apoB together with the amount of phenotypic variance explained by age, sex and BMI are given in Table 2.2.6.

Table 2.2.6 Mean serum lipid and lipoprotein traits and phenotypic variance explained by age, sex and BMI in the Swedish patient group.

	Mean (± SD) serum level	Phenotypic variance explained by differences in age, sex and BMI
Total Cholesterol mmol/l	7.17 ± 1.32	3.8%
Total Triglyceride mmol/l	2.13 ± 1.70	4.9%
LDL Cholesterol mmol/l	4.81 ± 1.14	5.7%
HDL Cholesterol mmol/l	1.15 ± 1.26	25%
ApoB mg/100ml	124.8 ± 22.3	6.1%

2.2.7 Chinese

A group of 83 healthy Chinese males, mean age 36 years, living in Singapore. A detailed description of this sample including apoB genotype data is presented in Saha et al. (1992a).

2.2.8 Afro-Caribbeans

A group of 47 healthy individuals aged 45 to 74 years (mean age 55 years), with at least three grandparents of Afro-Caribbean origin, drawn from family practitioner registers in two North-West London health centres. This sample has been described in detail in Miller et al. (1989).

2.2.9 South Asians

A random subset of 152 men drawn from a sample of 1433 South Asian men aged 40-65, living in the west of London, who have been investigated for CAD risk factors. The original sample was assembled from four factories and lists of general practitioners in west London. 89% of the men were Punjabi Sikhs and the rest were evenly divided between Gujarati and Punjabi Hindus. (McKeigue et al. 1991). ApoB genotyping is given in Renges et al. (1991).

2.2.10 St Thomas' Atherosclerosis Regression Study (STARS)

A group of 70 British patients, under 66 years of age with CAD as determined by coronary angiography. Patients included in the study had untreated plasma cholesterol levels between 6.0 and 10.0mmol/l and triglycerides below 4mmol/l. Patients who had suffered a myocardial infarction in the previous eight weeks, or who suffered from diabetes, hypertension, cardiac failure or malignancy were excluded from the study. After baseline assessment, patients were assigned into one of three treatment groups: lipid-lowering diet alone, cholestyramine plus the lipid-lowering diet or no treatment and were followed for three years before repeat angiography (Watts et al. 1992). ApoB genotyping is given in Peacock et al. (1994).

2.2.11 Camberley

A group of 364 caucasian men attending a general practitioner's surgery in Surrey, UK . All were aged between 50 and 61 years and had no clinical evidence of CAD (Humphries et al. 1993).

2.2.12 ECTIM Strasbourg patients.

A group of 211 men aged between 25 and 64 years, who had survived a myocardial infarction three to nine months prior to inclusion in the study. Their families had been resident in the Bas-Rhin department of north-east France for the previous two generations and all four grandparents had been born in Europe (Moreel et al. 1992).

2.2.13 ECTIM Strasbourg controls.

A group of 194 men, age-matched with the patient group drawn from the Strasbourg electoral rolls (Moreel et al. 1992)

2.3 DNA FROM HIGHER APES

DNA from 4 unrelated Chimpanzees, 4 unrelated Gorillas and 2 Orang-utans was also investigated in Chapter 4. The ape DNA was a kind gift from Dr I Gray and Prof. A Jeffreys.

2.4 Polymerase Chain Reaction (PCR)

2.4.1 Reactions using Taq Polymerase.

Unless otherwise stated regions of the apoB gene were amplified by the following method of PCR adapted from Saiki et al. (1988). Reactions consisted of approximately 50ng genomic DNA as template, 250ng each primer and 1unit *Thermus aquaticus* DNA polymerase (Perkin Elmer/Cetus CT. USA. and Gibco-BRL, Paisley, UK) in a buffer consisting of: 50mM KCl, 10mM Tris/HCL (pH8.3), 0.001% Gelatin, 0.1% W-1 detergent, 0.2mM each dNTP (Pharmacia, Uppsala, Sweden) and 0.5 to 4.0mM $MgCl_2$ (depending upon the primers used). After initial denaturing of the DNA for 5mins at 95°C, annealing of the primers (55°C, 3mins) and extension (72°C, 3mins) the program used was 35 cycles of: 98°C, 1min; 55°C, 1min and 72°C, 1 to 3 mins. Reactions were performed in automated thermal cyclers (Cambio, Cambridge, UK and Genetic Research Instrumentation Ltd, Essex UK).

2.4.2 Reactions using Vent polymerase.

This was used in the work described in Chapter 7 to reduce the number of errors introduced during the reaction. 500ng of each primer was mixed with 350ng of target genomic DNA in buffer supplied by the manufacturer with the addition of 400μM each dNTP and 4 to 6mM $MgSO_4$. Five units of Vent polymerase (New England Biolabs. Beverly MA, USA) were added and the reaction performed using a "touch-down" procedure to minimise false priming of the oligos. After an initial DNA denaturation at 95°C for 5 mins and primer annealing/extension at 75°C for 5 mins, the cycle was changed to 98°C denaturation for 1 min and 72°C extension for 2 mins. The 1 min annealing temperature was lowered by 5°C increments with each successive round of amplification until 60°C was reached. Thereafter, 35 identical cycles of 98°C 1min, 60°C 1min and 72°C 2mins, were carried out followed by a final extension period of 72°C for 15 mins.

2.5 Restriction Enzyme Digestion.

PCR amplified or plasmid DNA was digested with restriction enzymes according to the manufacturer's instructions in the appropriate buffer, either supplied by the manufacturer or from a five buffer set (Boehringer-Mannheim, UK) with the addition of 100ng/μl BSA (Gibco-BRL, Paisley UK). Digestions were carried out in water baths at the recommended temperature for the enzyme. For RFLP analysis, part or all of PCR reaction was normally digested with 10units of enzyme for 12 to 24 hours. Prior to cloning, 1 to 5 μg of PCR amplified or plasmid DNA was digested with 1 to 5 units of enzyme for one hour. If digestion was then found to be incomplete incubation would be continued for a further hour.

Restricted DNA was normally analysed by electrophoresis in agarose gels (Gibco-BRL, Paisley, UK) containing 1x TAE buffer. Gels were stained with 0.5μg/ml Ethidium Bromide prior to visualisation on a UV (260nm) transilluminator .

2.6 Purification of DNA.

If purification of amplified or cloned DNA was required one of the following methods was used: Phenol extraction and ethanol precipitation following the procedure given in Maniatis et al. (1982), or alternatively by using Geneclean (Bio101, La Jolla CA, USA) according to the manufacturers instructions. One of these methods was routinely used prior to sequencing amplified DNA and after restriction enzyme digestion during cloning operations.

2.7 Chemical Cleavage Mismatch Analysis

Base differences in target DNA were detected by a modification of the technique described by Cotton et al. (1988), (Montandon et al.,1989). PCR was performed on all target DNA, plus the negative control "probe" sample. Approximately 100ng of purified, amplified DNA from the "probe" individual was then end-labelled with 10pM [γ^{32}P] ATP (<5000Ci/ mmol, Amersham, UK) using T4 Polynucleotide Kinase (Gibco-BRL, Paisley, UK) and the labelled products were purified by centrifugation through a Sephadex

G-25 (Pharmacia, Uppsala, Sweden) column in a volume of 100μl 10mM Tris (pH 7.6) / 1mM Na_2 EDTA (TE) buffer.

Heteroduplex DNA was formed by mixing 10ng of labelled "probe" to 100ng purified, amplified target DNA in 50μl of 0.3M NaCl/0.1M Tris (pH8.0), boiling for 5 mins then incubating for 1 hour at 65°C. Duplex DNA was ethanol precipitated overnight at -20°C with the addition of 40μg mussel glycogen (Boehringer-Mannheim, UK) as a carrier. The resulting pellet was washed once with ice-cold 70% ethanol and dried briefly before resuspension in 16μl TE buffer. 7μl of the heteroduplex solution was made up to 20μl by the addition 4mM hydroxylamine hydrochloride solution (Aldrich, UK) [titrated to pH 6.0 with diethylamine (Aldrich)]. This was incubated at 37°C for 2 hours. The remaining heteroduplex DNA was added to 16μl of a buffer consisting of 5mM Tris (pH 8.0), 0.5mM EDTA, 3% pyridine and 0.025% osmium tetroxide (Aldrich, UK) and incubated at 37°C for 2 hours. DNA from both reactions was ethanol precipitated overnight, and after washing was resuspended in 50μl 1M piperidine (Aldrich) and incubated at 90°C for 45mins. After a final ethanol precipitation, samples were resuspended in 2μl stop solution (95% formamide/4% EDTA/0.5% bromophenol blue/0.5% xylene cyanol) and run on denaturing polyacrylamide gels as for sequencing (Section2.8).

2.8 Sequencing.

2.8.1 Sequencing of cloned DNA.

Phage M13 mp11 (Pharmacia, Uppsala, Sweden) clones containing inserts were identified by comparison with native M13 on 1% agarose gels - clones containing inserts larger than 300bp migrate more slowly. Such clones were purified ready for sequencing using T7 DNA polymerase (Sequenase, USB, Cleveland USA) following the manufacturer's recommended protocol. Reactions were primed with M13 universal forward or reverse primers (Sequenase, USB, Cleveland USA). Internal primers, determined from the apoB gene sequence (Cladaras et al. 1986) and manufactured to order, were used to resolve sequences furthest from the cloning sites. Sequencing reactions were resolved on 8% polyacrylamide gels (19:1 acrylamide to bis-acrylamide ratio) (Severn Biotech Ltd., Kidderminster, UK) with the addition of 7M urea as denaturant in 1xTBE buffer, on a BRL model S2 apparatus (Gibco-BRL, Paisley UK). After

running, gels were dried and autoradiographed for 12 to 96 hours on Hyperfilmβ max (Amersham, UK).

2.8.2 Direct Sequencing of Amplified DNA - 1st Method

This method is based on that of Green et al. (1989) and used modified T7 DNA polymerase (Sequenase, USB, Cleveland USA). DNA was amplified by PCR, purified by Geneclean and resuspended in TE to a concentration of approximately 100ng/μl. Annealing was carried out in a total volume of 7μl with 200ng amplified fragment (approx. 0.5pmol) and 20pmol oligonucleotide primer in Sequenase buffer containing 0.5%NP-40 (Sigma, UK). The annealing mixture was boiled for 10 mins then snap-cooled on dry-ice. A "labelling" mixture for 10 reactions was made, containing 50μCi $\alpha[^{35}S]$dATP (Amersham, UK) 10μl 0.1M dithiothreitol, 4.5μl 10%NP-40 and 1.5μl Sequenase enzyme in TE buffer to a total volume of 45μl. 4μl of this mixture were added to each annealed primer/template. 2μl of this new mixture were then immediately added to each of four tubes containing 2μl of "A","C","G" and "T" termination mixes respectively. The termination mixes contained 80μM each of dCTP, dGTP, and dTTP and 8μM of the appropriate ddNTP, except for the "A" mix which contained 80μM of dCTP, dGTP, and dTTP and 0.08μM ddATP. After 5 mins of incubation at 37^{0}C, reactions were chased with 2μl of a solution containing 0.25μM each dNTP and 0.05%NP-40 and incubated for a further 5 mins at 37^{0}C. Reactions were stopped with the addition of 4μl of stop solution (95% formamide/4% EDTA/0.5% bromophenol blue/0.5% xylene cyanol). 4μl of the final volume were boiled for 10 mins then immediately loaded onto gels (as Section 2.8.1).

2.8.3 Direct Sequencing of Amplified DNA - Improved Method.

An alternative method of directly sequencing PCR amplified DNA utilised streptavidin coated magnetic beads (Dynal AS, Norway). PCR was carried out using one normal primer and the other labelled with biotin at the 5'end. 40μl of PCR product were then mixed with an equal volume of Dynabeads (Dynal AS, Norway) in a solution of 10mM Tris-HCl (pH7.5), 1mM EDTA and 2.0M NaCl and incubated at room temperature for 15 mins. The supernatant was then removed whilst holding the beads in the tube with a magnet (Dynal AS, Norway). The beads were washed once in 40μl of 10mM Tris-HCl (pH7.5)/ 1mM EDTA /1.0M NaCl and then resuspended in 8μl 0.1M NaOH and incubated for 10 mins. After removal

of this denaturing solution, the beads were washed successively with 50μl each of: 0.1M NaOH, 10mM Tris-HCl (pH7.5)/ 1mM EDTA /1.0M NaCl and TE buffer. After removal of the final wash, beads (attached to the single-strand of PCR DNA with the biotin labelled primer at the 5' end) were resuspended in 7μl TE and sequenced using T7 DNA polymerase (Sequenase, USB, Cleveland USA) following the manufacturer's protocol. Before loading the gels, reactions were boiled for 10 mins and then the beads (still attached to the template DNA strand) were held in the magnet whilst loading 4μl of the supernatant (containing the terminated strands). Gel running, drying and autoradiography was as in Section 2.8.1.

2.9 Screening for Single Base Changes in amplified DNA

2.9.1 Blotting and Hybridisation to Allele Specific Oligonucleotides [ASOs]

PCR reactions were performed and visualised on 0.7-1% agarose gels. Gels were denatured for 30 - 60mins in a solution of 1.5MNaCl/0.5MNaOH and then sandwich-blotted onto nylon membranes (Hybond N+, Amersham, UK) using the denaturing solution in the gel as transfer buffer. After blotting for 2 to 16 hours, blots were neutralised in a solution of 2xSSC/0.1MTris-HCl(pH7.5) for 5 mins before baking at 80^{0}C for 2 hours to fix the DNA onto the blot. ASOs were labelled at the 5' end with T4 Polynucleotide Kinase (BRL, Paisley, UK) and [γ^{32}P] ATP (Amersham, UK) to a specific activity of approximately 0.1μCi/pmol (Maniatis et al. 1982). Filters were hybridised for 3 -16 hours in 5x SSPE/ 0.5% SDS/ 5x Denhardt's solution - 0.1mls of hybridisation were used with approximately 0.5ng of one of the labelled ASO's per 1cm^2 of nylon membrane. Hybridisation was performed in tubes in a rotating oven (Bachoffer GmBH, Germany) at 5^{0}C below the melting temperature for the oligonucleotide, and then washed in excess 5x SSPE/ 0.1% SDS for 10mins at the melting temperature of the ASO. Autoradiography was carried out for 2 to 16 hours at -70^{0}C on X-ray film (Fuji Ltd, UK) using intensifying screens.

2.9.2 Non-radioactive adaptation of ASO method.

The results in Chapter 6 were obtained using oligonucleotides labelled at the 5' end with Fluorescein-dUTP by Terminal Transferase (Amersham, UK) using the manufacturer's recommended protocol. Blotting and hybridisation was carried out as above but hybridisation was in 5xSSC/0.1%SDS/0.5% Blocking agent (Amersham, UK) containing 10ng probe/ml solution, followed by washing twice in excess

1xSCC/0.1%SDS at the melting temperature of the oligonucleotide. ASOs were then detected using an anti-Fluorescein antibody conjugated to horse-radish-peroxidase (Amersham, UK) in the Enhanced Chemi-Luminescent (ECL) system (Amersham,UK) following the manufacturer's protocol. Light detection was on X-ray film (Fuji, UK) for 3-120mins.

2.9.3 Detection of Restriction Fragment Length Polymorphisms [RFLPs]

Thr_{71} to Ile [Ag(g/c)] was detected as an ApaLI RFLP (Ma et al. 1989). Exon 4 of the apoB gene was amplified using oligonucleotides situated in introns 3 and 4 (Sequences: 5' oligo GATAAGGCATGTGGTGTTGAGGCGC, 3' oligo TCCGTGCCTGGTGCAAACACACAAG - both given in 5' to 3' orientation) to give a 373bp fragment. 10% of the amplified DNA was digested with 10 units ApaLI (Bioexcellance, Cambridge, UK) The enzyme cuts the sequence GTGCAC, which encodes Thr_{71}, to give two fragments of 200bp and 173bp. The C to T base change which generates the codon for Ile_{71} also destroys the ApaLI recognition site leaving an undigested 373bp amplified fragment. The DNA fragments were visualised on 2% agarose gels.

Val_{591} to Ala [Ag(a/d)] was detected as an Alu I RFLP (Wang et al. 1988). Oligo primers (Sequences: 5' GGAAATCCATATTTACTTGGGGTGC and 3' CTTCTATTTTGGCTGAGGCTGGGTC) were used to amplify 82bp of intron 13 together with the 5' end of exon 14, generating a 223 bp fragment. 10% of the PCR was digested with 10units AluI to give either two fragments of 60bp and 167bp (in alleles encoding Val_{591},) or three fragments of 60bp, 50bp and 117bp (in alleles encoding Ala_{591}). DNA fragments were visualised as above.

The XbaI RFLP was detected in the higher apes after amplification of the DNA with primers (Sequences: 5' oligo GGAGACTATTCAGAAGCTAA and 3' oligo CAAGAGCCTGAAGACTGACT) to give a 700bp fragment. 10% of the PCR reaction was digested with 10units XbaI which recognises the site AGACTC (Bioexcellance, Cambridge, UK) to give two fragments of sizes 425bp and 275bp in X+ alleles. The DNA fragments were visualised as above.

2.10 Cloning of DNA

2.10.1 Ligation

Insert DNA, generated by PCR, and plasmid or phage M13 DNA were digested with appropriate restriction enzymes, purified by phenol extraction or Geneclean and resuspended in TE buffer. An approximately 3:1 molar ratio of insert to vector DNA (using 100ng of vector DNA) was ligated using 10units T4 DNA Ligase (Promega, Madison WI, USA) in a volume of 10μl containing the manufacturer's supplied buffer at 14^0 C for 48 hours.

2.10.2 Transformation

Ligations were diluted 10 fold and used to transform competent *E. coli* strains JM101 (for phage M13), XL1-Blue (Strategene, La Jolla CA, USA) or DH5α (Promega, Madison WI, USA) (for generating expression vectors). 5μl diluted ligations were mixed with 100μl of competent cells. These were incubated for 30 mins on ice, heat shocked at 42^{0}C for 2 mins, allowed to recover in 1ml of Luria broth at 37^{0}C for 1 hour and then spread between 2 plates containing Luria agar with the addition of 160μg/ml IPTG and 40μg/ml X-gal (Promega, Madison WI, USA). In the transformation of plasmids 50μg/ml ampicillin and/or 12.5μg/ml tetracycline were also added, depending on the resistances conferred by the different vectors. Plates were incubated at 37^{0}C for 16 hours and then 2 hours at 4^{0}C to allow for blue colour development.

White colonies were picked and grown for 16 hours in 3mls of Luria broth containing appropriate antibiotics to maintain resistance. A secondary screening for clones containing apoB exon 29 sequence (chapter 6) was done using PCR. 2μl of the Luria broth were heated to 95°C for 5 mins to release and denature the DNA and then amplified using primers internal to exon 29 which together cover nucleotides 12585 to 13379 of the apoB gene :-5' Primer - ^{5}'CCAGGAACTGTTGACTCAGGAAGGCCAAGCC and 3' Primer - ^{5}'GAGCAACTAACAGGTTCTTGATCAGAC.

2.10.3 Preparation and purification of plasmid and phage M13 DNA.

Plasmid DNA was prepared from 16 hour cultures (either 3ml, for mini-preps, or 250mls for large scale preparations) by the alkaline lysis method (Maniatis et al. 1982). Before further use DNA was

purified either by centrifugation to equilibrium in caesium chloride-ethidium bromide gradients (Maniatis et al. 1982) or by precipitation with NaCl/ polyethylene glycol (Assayable et al. 1990).

Single- stranded Phage M13 DNA was prepared from 5 hour incubations of one phage plaque into 1.5mls of log-pase *E. coli* (Strain K12, Amersham, UK). Incubations were carried out at 37°C with constant agitation. Cells were then pelleted in a micro-centrifuge and the supernatant retained. Phage were precipitated from the supernatant by the addition of 200μl of 20% polyethylene glycol 6000/ 2.5MNaCl and incubation for 15mins at room-temperature, followed by centrifugation in a micro-centrifuge. Viral pellets were resuspended in 100μl TE buffer, phenol extracted and ethanol precipitated before resuspension in a final volume of 50μl TE buffer.

2.11 Expression in Insect Cell Lines.

2.11.1 Culture of Sf-21 cells and infection with baculovirus

Insect Sf-21 cells were cultured in suspension in TC-100 medium containing 10% foetal calf serum (Gibco-BRL, Paisley UK), at 27°C with increased humidity but no CO_2, up to a density of $1x10^6$ cells/ml. Cells were then seeded at the required density onto petri dishes and allowed to adhere for 30 mins at room temperature before infection with virus. Media was removed from cells by aspiration, prior to infection by adding virus stock drop-wise onto the cells and incubating at room temperature with gentle rocking for one hour. Plates were then supplemented with 4 - 10 ml more TC-100 medium, containing 10% foetal calf serum, and incubated at 27°C with humidity for 4 days, unless otherwise stated. Viral plaques are seen under a high-power, phase-contrast microscope, as a group of larger, more opaque cells surrounding a region of lysed cells in the monolayer. Non-recombinant virus still contains an intact polyhedron gene and so cells infected with these are seen to contain bright granules of polyhedron protein. Recombinant virus has had the polyhedron gene disrupted by the apoB insert and so the cells infected with these do not contain the bright granules.

2.11.2 Cationic Liposome Mediated Transfection of Sf-21 Cells

3μg of pVT-Bac (Invitrogen, San Diego, CA, USA) containing exon 29 insert (purified by ultra-centrifugation through $CsCl_2$ gradient) was mixed with 1μg of linear AcMNVP DNA (Invitrogen, San Diego, CA, USA) in 1ml of TC-100 medium. This was then vortexed with 20μl of cationic liposome solution (Invitrogen, San Diego CA, USA) and incubated for 15 mins at room temperature. This mix was used to transfect $2x10^6$ Sf-21 cells, in a 60mm tissue culture petri dish, by rocking for 4 hours at room temp. After transfection a further 1ml of TC-100 medium supplemented with 10% foetal calf serum was added and the cells incubated for 60 hours. Virus stock was made by harvesting the medium covering the cells, centrifuging at 1500 rpm to pellet the cell debris and then transferring the supernatant to a fresh tube.

2.11.3 Plaque purification

$3x10^6$ Sf-21 cells were infected with 1ml each of either a 10^{-2} or 10^{-3} dilution of the virus stock. Cells were then overlaid with 10ml of TC -100/10% foetal calf serum/0.6% agarose and incubated for 4 days. Recombinant plaques were identified under a phase-contrast microscope. Well isolated plaques were picked with a pasteur pipette into 1ml TC -100 with 10% foetal calf serum.

2.11.4 Amplification of the viral titre

$2.5x10^6$ Sf-21 cells were inoculated with the 1ml purified plaque solution. A further 5mls of complete TC-100 then added. The medium covering the cells was harvested after 4 days and stored as Pass-1 virus stock, after centrifugation at 1500 rpm to pellet the cell debris. Pass-2 virus stock was made in the same manner except that $8x10^6$ Sf-21 cells were inoculated with the 0.5ml of the Pass-1 stock.

2.11.5 Time course peptide expression study

$4x10^6$ Sf-21 cells were infected with 1ml of the Pass-2 virus and 0.5ml of the Pass-1 virus for 1 hour. Cells were then washed 3 times with TC-100 medium to remove all foetal calf serum and finally covered with 3mls SF-900 serum free complete medium (Gibco-BRL, Paisley, UK). At 12 hour time points subsequent to infection, 200μl alliquots were removed and clarified by centrifugation at 1500rpm. Protease inhibitors: benzamidine at 3mg/ml and Phenylmethylsulphonylfluoride at 0.2mg/ml were added to the supernatant before storage at 4^0C.

Maximum Likelihood procedure[#]

Initially the numbers of unambiguous haplotypes were calculated and then the haplotypes of the doubly heterozygous individuals were assumed to be equally distributed between A1B1/A2B2 and A1B2/A2/B1.

The probability of the double heterozygotes being distributed as A1B1/A2B2:-

$$= \frac{\text{number of A1B1/A2B2 haplotypes}}{\text{total number of haplotypes observed}(n)}$$

$$= \frac{a/n \times d/n}{[a/n \times d/n] + [b/n \times c/n]} = P_1$$

The value P_1 was then used to redistribute the haplotypes of the doubly heterozygous indivdiuals in the proportions :-

P_1 (A1B1/A2B2) and $(1-P_1)$ (A1B2/A1B2)

The value P_2 was then calculated using the new distribution of haplotypes and this calculation was reiterated until the values of P no longer changed (A. Templeton, personal communication).

2.11.6 Western Blotting

Peptide samples were separated on 10% polyacrylamide/SDS gels and electro-blotted onto nitrocellulose membranes (Bio-rad, Hemel Hempstead, UK). Blots were blocked for 2 to 16 hours in 1xTBST [25mM Tris(pH7.5), 0.15M NaCl, 0.05% Tween-20]/2% non-fat dried milk/0.01% sodium azide and then incubated with first antibody at a 1/10000 dilution in more of the same solution for 2 to 5 hours at room temperature. After washing three times in excess 1xTBST/0.5% non-fat dried milk, protein was detected using goat anti-mouse antibody and the ECL system (Amersham,UK) according to manufacturer's instructions.

2.12 Statistical Analysis

2.12.1 Gene frequencies and Hardy-Weinberg equilibrium.

The gene-counting method was used for the estimation of gene frequencies, 95% confidence intervals for sample allele frequencies were calculated according to Colton (1974) = $p \pm 1.96\sqrt{p(1-p)/n}$. Where p = proportion of study sample with given allele and n = no. of participants.

Comparison of allele frequencies between population samples was done by X^2 analysis. Estimation of Hardy-Weinberg equilibrium was done by comparing the expected genotype distribution for a given allele frequency ($p^2 + 2pq + q^2 = 1.0$)* with the observed distribution using a X^2 test. Statistical significance was considered to be at the 0.05 level.

* where p the frequency of one allele and q is the frequency of the other and $p+q=1.0$.

2.12.2 Detection of Linkage Disequilibrium between alleles at different polymorphic sites.

Linkage disequilibrium between pairs of allelic sites was estimated using the correlation coefficient Δ (Chakravarti et al. 1984)

$$\Delta = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

where a,b.c and d are the calculated numbers of the haplotypes A1B1, A1B2, A2B1, and A2B2 respectively. The haplotypes of the doubly heterozygous individuals were estimated using a maximum likelihood procedure. 95% confidence intervals were calculated using the z transformation, z has a normal

distribution with a standard error of $1.96/\sqrt{n-3}$.

2.12.3 Calculation of percentage of variance explained by genotypes.

The percentage of variance (R^2 x 100) explained by different factors in the Swedish patient and control groups was estimated by multiple stepwise regression using the SPSS/PC+ computer program. The skewness of the total cholesterol, triglyceride, LDL-cholesterol and HDL-cholesterol distributions in the controls and of the triglyceride and HDL-cholesterol distributions in the patients was adjusted by $\log_{10}$ transformation prior to analysis, all other traits investigated were normally distributed. After adjustment for age and BMI in the control group and age, sex and BMI in the patient group, the impact of the apo B genotypes on the lipid and lipoprotein traits were estimated by linear regression. Genotype classes were re-coded to avoid biasing the results. Statistical significance was considered to be at the 0.05 level.

2.12.4 Comparison of differences between group means.

For each adjusted variable, a one-way analysis of variance (ANOVA) was performed to compare differences between group means according to genotype or haplotype class.

2.12.5 Analysis of apoB levels by centiles within a population sample.

Lipoprotein or apo B levels were expressed as centiles separately for each of the samples examined. Individuals carrying particular haplotypes or variants were then identified within the different centile. divisions.

2.13 Secondary Structure Protein Modelling

The protein secondary structure was modelled using the one or both of the following programs: The "Predict" suite of ten secondary structure prediction programs (Dept. of Biophysics, University of Leeds, UK) - secondary structure predicted by this program is approximately 60% reliable. The DNA-star "Protein" program estimates secondary structure by the Chou-Fassman and Garnier-Robson algorithms. Hydrophobicity is estimated from the Hopp-Woods algorithm. Charge and hydrophobic moment (a measure

of amphipathicity) are also calculated. A window of seven residues was used in the estimations. Segments of amino acid sequence extending 250 residues either side of the polymorphic sites were analysed. The results from both programs are in general agreement for the polypeptide encoded by exon 29.

MATERIALS

2.14 Suppliers of Materials and Reagents

Oligonucleotides were either synthesised on a Pharmacia Gene Assembler (Pharmacia, Uppsala, Sweden), or they were bought from Severn Biotech Ltd. (Kidderminster, UK.), Oswell Scientific Ltd, (Edinburgh, UK.), Advanced Biotechnology Centre (Charing Cross and Westminster Hospital Medical School, London, UK.) and Genosys Ltd (Cambridge,UK). Unless otherwise stated, all chemicals were obtained from BDH Ltd or Sigma Ltd, Poole, Dorset, UK. Suppliers of enzymes, kits and radionucleotides are given in the text.

The constituents of standard solutions and culture media referred to in this thesis: 20xSSC, 20xSSPE, 10xTAE, 10xTBE, TE and Luria broth, are given in Maniatis et al. (1982).

RESULTS AND DISCUSSION

CHAPTER 3

THE SEARCH FOR MUTATIONS IN THE PRIMARY RECEPTOR BINDING DOMAIN OF APOB [RESIDUES 3130 TO 3630].

Summary

Sequence differences in the region of the apoB gene encoding amino acids 3130 - 3630 were sought in DNA from eight individuals with reduced affinity of Low Density Lipoprotein (LDL) for the normal LDL-receptor. All individuals were hypercholesterolaemic and were selected either on the basis of reduced Fractional Catabolic Rate (FCR) of autologous LDL, or substantially reduced binding of their LDL to normal LDL-receptors determined by an *in vitro* cell growth assay using the U937 macrophage-like cell line. Segments of the apoB gene were amplified by the polymerase chain reaction. Using a combination of cloning and sequencing the amplified fragment, together with chemical cleavage mismatch analysis, no sequence differences were identified in this region of the gene. It was therefore concluded that variation outside the region of the apoB gene that codes for amino acids 3130-3630 must be responsible for the reduced LDL clearance in these patients.

3.1 Rationale for study.

Two studies have been performed to investigate the association between the apoB XbaI genotype and *in vivo* turnover of autologous LDL (Demant et al. 1988, Houlston et al. 1988, Section 1.5.4) and both showed an association of the X+ allele with reduced FCR of LDL. This indicated that the raised serum apoB, LDL and total cholesterol levels, also associated with the X+ allele, might be mediated through reduced clearance of LDL via the LDL receptor, rather than through overproduction of LDL. Thus the search for a lipid raising mutation in allelic

association with the XbaI X+ allele was initially centred on the region of the gene encoding the apoB primary receptor binding domain (residues 3130 -3630 - Section 1.1.1).

3.2 Individuals studied

All of the ten patients studied in this chapter have been shown to be hypercholesterolaemic on several occasions, even after dietary advice. A description of them is given in Section 2.2.1. Biochemical and apoB genotype details of these patients are shown in Table 2.2.1. Patients 1 to 7 and 9 have had LDL FCR determined by *in vivo* turnover studies with autologous LDL (Shepherd et al. 1979, Turner et al. 1984, International Collaborative Study Group, 1986). Patients 6, 7, 8 and 10 have demonstrated significantly reduced growth propagation in the U937 Cell Growth Assay (Frostegard et al. 1990) with two or three plasma samples taken several months apart.

3.3 Gene region and methods of study

The 1.54kb region of exon 26 of the apoB gene studied in this chapter is shown in Figure 3.3.1. This region includes the two sequences encoding protein regions with homology to the apoE receptor binding domain (nucleotides 9648 - 9681 and 10245 - 10353), and the sites encoding the Arg_{3500} to Gln (nucleotide 10708) and Arg_{3611} to Gln (nucleotide 11041) substitutions. The entire region was amplified in two overlapping sections using PCR. The positions and sequences of the primers are given in Figure 3.3.1. Primers A1/A2 generate a 5' fragment 798bp long, and primers B3/B4 yield a 3' fragment 788bp long. The A1/A2 PCR products from each patient were digested with restriction enzymes ScaI and EcoRI and ligated into complementary sites in phage M13 mp11 using T4 DNA ligase. Competent *E.Coli* strain JM101 were transformed with these phage and plated out for blue/white selection.

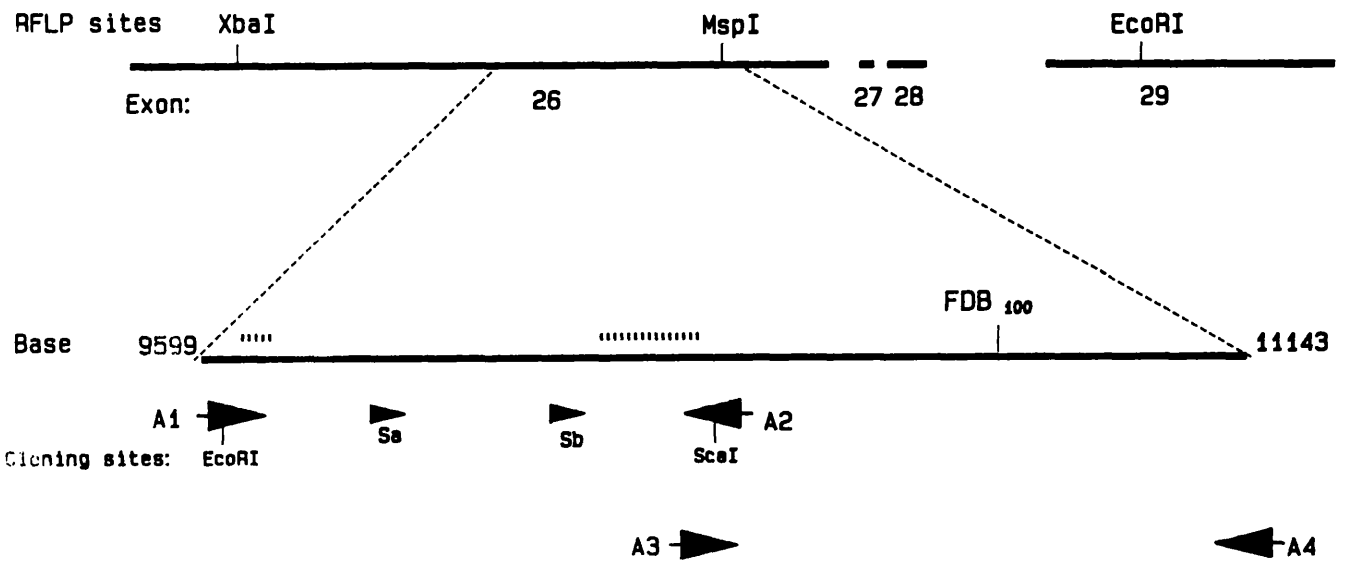


Figure 3.3.1 Map of the 3' end of the apoB gene

The positions of oligonucleotides used for PCR and sequencing are shown. The bold dotted lines indicate the nucleotides encoding protein regions with homology to the apoE receptor binding domain (nucleotides 9648-9681 and 10245-10353). The relative positions of base changes creating RFLP sites and the $apoB_{3500}$ mutation are shown. Primer sequences are: (5' to 3')

A1 - TTGAAGGAATTCTTGAAAACGACAAAGCAA, A2 - CACTTCCATATTTTTCGTGGTTAAGCTCAC, A3 - CATAACAGTACTGTGAGCTTAACCACGAAA,

A4 - AAGGATCC TGCAATGTCAAGGTGTGCCTTT, Sa - CAAA TTCCTGGATACACTGT, Sb - GAATACCAATGCTGAACTTT.

Single white plaques were picked, purified and sequenced (as described in Section 2.8.1). The sequences were resolved using the M13 Universal primer and internal primers Sa and Sb (Figure 3.3.1). Initially only one clone from each patient was examined, but later 10 clones from each patient were sequenced (sections 3.4 and 3.5). In order to detect any base changes rapidly, the 10 clones sequenced were run with all the tracks of the same base together.

Both amplified regions, A1/A2 and B3/B4, were also examined by Chemical Cleavage Mismatch Analysis. DNA from patient 9 served as the "probe" for all the other samples and as the negative control. DNA from patient 10, heterozygous for G and A at residues 10708 and 11041, served as a positive control for the region B3/B4.

3.4 Sequencing the 5'PCR product A1/A2.

After cloning the products of this PCR from each of the 10 study individuals, the sequence of one clone from each individual was initially determined. It was assumed that this would give the sequence of one allele from each patient. Thus, the intention was to determine the sequence of at least seven X+ alleles (from patients 1 to 7) and one X- allele (from patient 9) and in addition either an X+ or X- allele would be sequenced from heterozygous patients 8 and 10. This strategy would mean that any base change, in linkage disequilibrium with the presence of the XbaI cutting site and occurring in the region studied, would be seen in more than one of the X+ alleles examined.

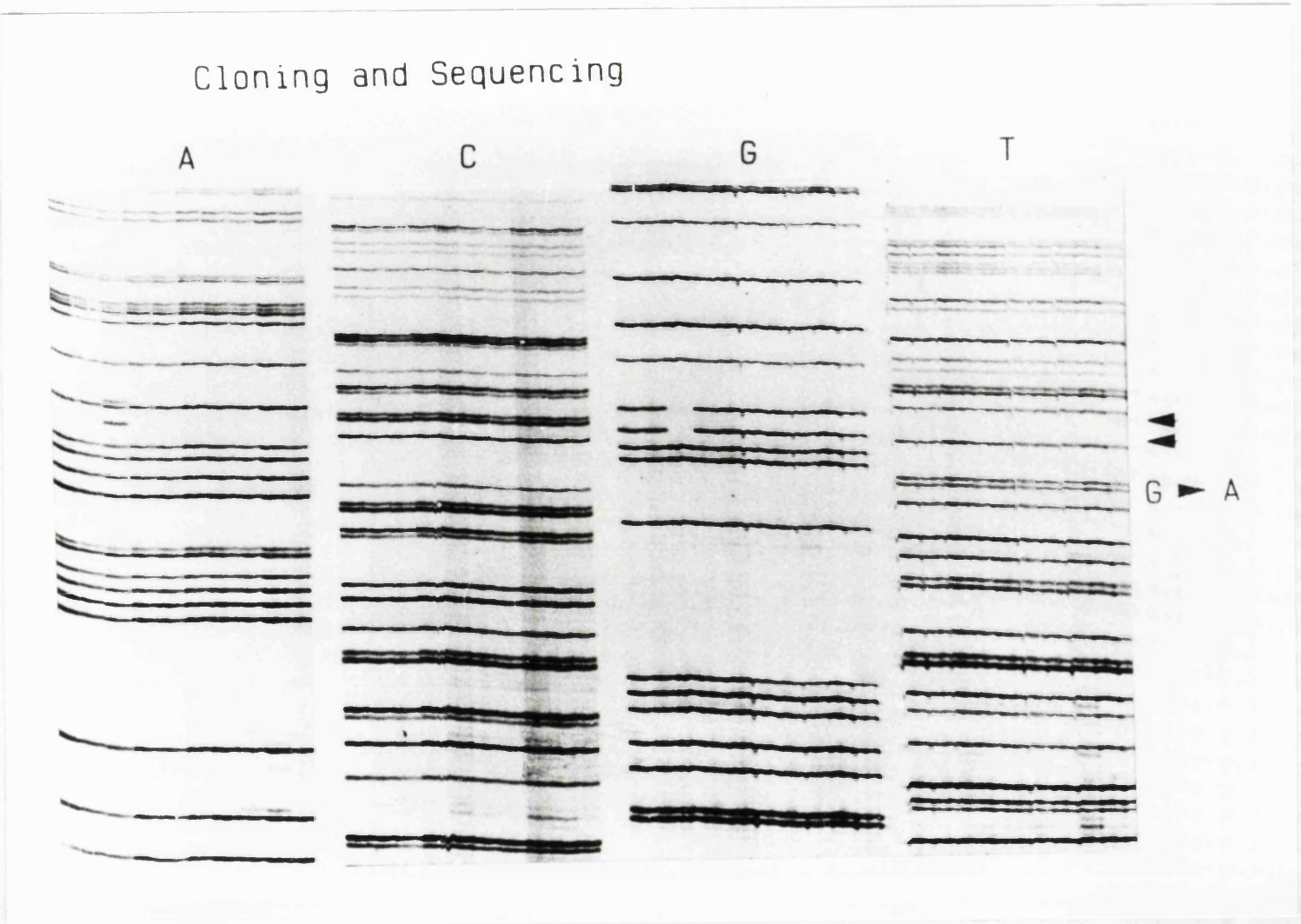


Figure 3.4.1 Example of cloning and sequencing.

Ten clones derived from patient 8, sequenced by priming with oligonucleotide Sa. The sequence of nucleotides 9981 to 10100 are shown. Two G to A mis-incorporation errors generated during the PCR reaction can be seen in clone 3 - indicated by arrows. The order of bases is given along the top.

When the results were obtained, 22 different base changes were seen in the first 10 clones sequenced (Table 3.5.1). No clone was identical to the published sequence (Olofsson et al. 1987) and none of the individual base changes were seen in more than one of the clones. Next, 10 clones derived from each patient were sequenced, since any base change present in the patient's genomic DNA in the heterozygous state, would be present in approximately half the clones derived from that patient (or all of them if the individual were homozygous for the change). None of the initial base differences were seen in any of the other clones (Figure 3.4.1) although all of them were reproduced upon re-sequencing the original clone. This indicates that all these base differences were artifacts generated by errors during PCR, and subsequently cloned.

Since no base change was seen in more than one clone, it can be concluded that any real base change in linkage disequilibrium with the XbaI site is not present in the 798bp region of exon 26 sequenced. Thus the mutation that raises serum apoB and cholesterol levels, being sought, was not present in this region of the gene.

3.5 Errors in the polymerase chain reaction.

The 22 PCR errors found whilst sequencing the first 10 clones (equivalent to 8000 bases of sequence determined) are given in Table 3.5.1. Also given are the sequences of the nucleotides immediately adjacent to each change. The most common base-substitutions found were A to G (5/22) and T to C (6/22). Three quarters of the changes seen were associated with a run of the same base. In three instances a base was deleted rather than substituted, however no insertions were observed in this experiment.

Table 3.5.1 PCR errors found

Base Change	Local Sequence
A to C	ATTTA**A**GT
A to G	TCACA**A**AT CATCA**A**AT TACA**A**AGC GATTG**A**AG GCCAC**A**GC
C to T	CTCCC**C**AG AGGCA**C**AG
G to A	GAGCT**G**CC TGCCA**G**TC
T to A	AGTTGTCA CAGCATGC TGCACTGC
T to C	GTTTATCA GGACCTTT ATGATTTC ATTGTTGC CTTCATTG TGAAGTTA
deleted A deleted G deleted T	GAAAA**A**GG TTCCA**G**TT AAGTTTGA

The nature of the errors made during PCR may give some insights into how they are generated. Taq Polymerase has no proof reading activity and so errors made cannot be corrected,

they are replicated on each subsequent round of amplification. The high incidence of errors associated with runs of the same base suggests that the enzyme may "stutter" over such runs, either omitting the base or inserting an inappropriate one. In 7/19 cases where a mispaired base was added, this base was identical to the base immediately preceding or following the substituted base. One possibility is that at the high temperatures at which PCR is performed (55-95°C) the DNA is always partially or completely melted and so slippage between the template and nascent strands may occur.

The frequency of the errors seen indicates that although this technique is sufficiently sensitive to detect single base changes, the cloning of DNA amplified with Taq Polymerase will tend to generate artifactual base substitutions and so sequences determined in this manner must be interpreted with caution. Paabo and Wilson (1988) suggested that the direct sequencing of PCR products (omitting the cloning step) would overcome this problem. Providing that sufficient molecules of template genomic DNA are put into the reaction, it is extremely unlikely that any error generated during the PCR will be amplified in excess of the genomic sequence. Thus, although many of the amplified DNA strands will carry errors at some point along their length, these will be competed out, during the sequencing reaction, by the excess of molecules present with the normal sequence at any given position. This approach has been used in subsequent experiments.

3.6 Screening the entire region by chemical cleavage mismatch analysis.

Another rapid screening technique which also overcomes the problem of cloning artifactual base changes is chemical cleavage mismatch analysis - a technique described by Cotton et al. (1988) and modified by Montandon et al. (1989) (Section 2.7). Briefly, heteroduplex DNA is formed between the labelled "probe" and unlabelled test DNA. This is then treated with either hydroxylamine, which will modify any unpaired C bases, or osmium tetroxide, which modifies

unpaired T bases. Piperidine is then used to break the DNA backbone at any modified bases. Finally the treated DNA is run on denaturing polyacrylamide gels. Only the labelled "probe" DNA is visualised. Uncleaved DNA is seen as a band the same length as the original PCR product. If, however, the test DNA has a base change and so does not fully protect the probe DNA, cleavage products corresponding in length to the position of the base change in the test DNA are seen.

Both amplified fragments : A1/A2 and B3/B4 were examined using this technique on DNA from all ten patients. Amplified DNA from patient 9 (homozygous X-X-) was labelled with ^{32}P ATP and used as the "probe" and negative control in both reactions. No cleavage products were seen when the PCR product of primers A1/A2 were investigated by this method (results not shown). This confirms the finding in Section 3.4 that no base changes exist in this gene region in the patients examined. Two positive controls were available for the B3/B4 region: Both patients 2 and 10 are heterozygous for the G to A base change at nucleotide 11041 which encodes Arg_{3611} to Gln. In addition patient 10 is heterozygous for the G to A substitution generating Arg_{3500} to Gln. Both these G to A substitutions are detectable as a mismatched C in the probe DNA upon modification by hydroxylamine. The results of this experiment in 6 of the patients studied are shown in Fig 3.6.1. An 86bp cleavage product can be seen in the hydroxylamine track of samples 2 and 10, an additional 412bp cleavage product can also be seen in patient 10. These correspond to the G to A substitutions generating the amino acid changes at residues 3611 and 3500 respectively. No other cleavage products were seen in any of the patients examined. Thus, it can be concluded that the chemical cleavage mismatch analysis technique is sufficiently sensitive to detect heterozygous base changes, but that other than in the positive controls described above, none were present in this gene region in these patients.

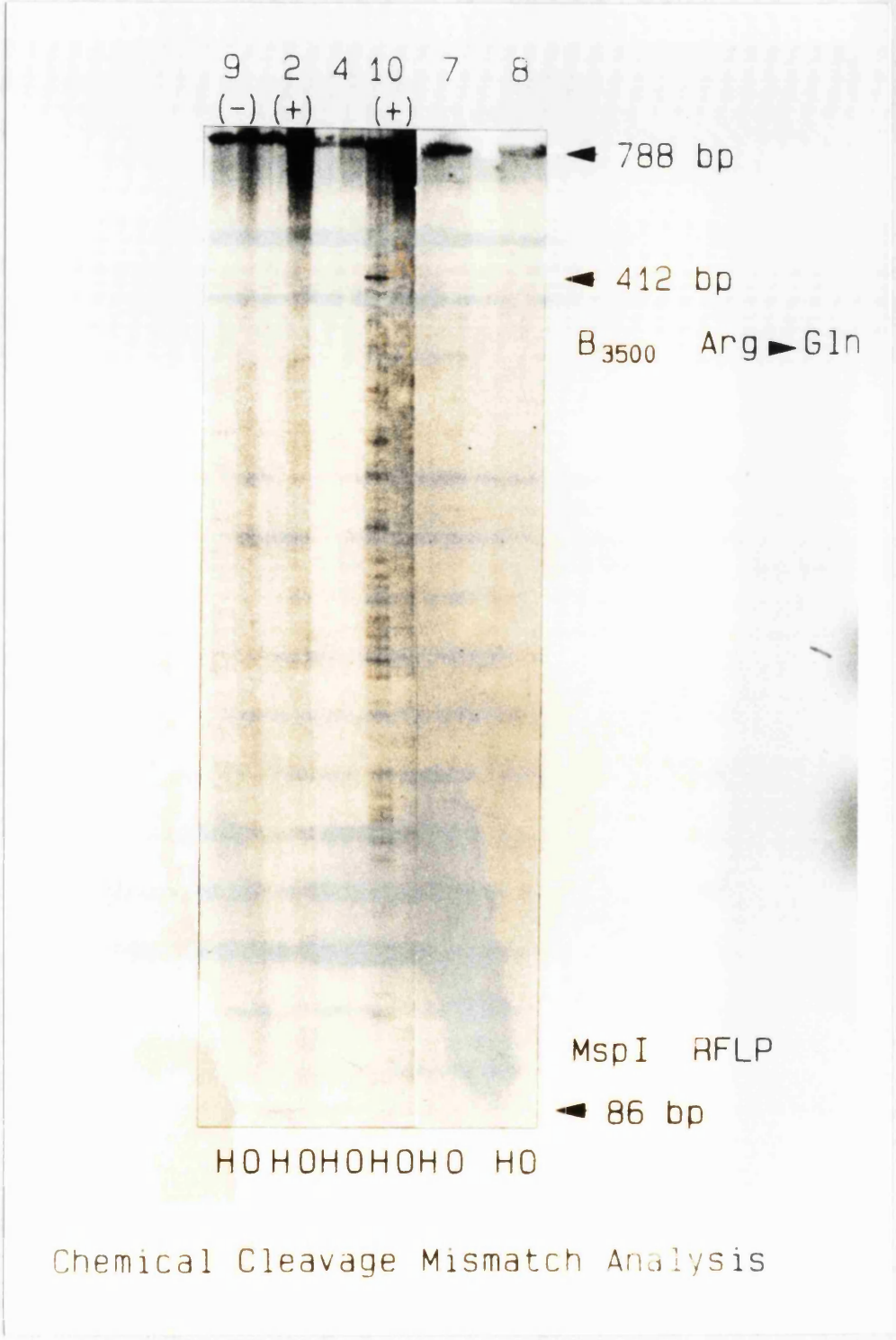


Figure 3.6.1 Example of chemical mismatch cleavage analysis of samples amplified with primers B3/B4.

Patients code numbers are given along the top. There are two tracks per patient: the hydroxylamine reaction (H), and the osmium tetroxide reaction (O). The uncleaved, full-length PCR gives a band of 788bp. DNA from patient 9 was used as the "probe" for all the other samples and as the negative control (-). DNA from patients 2 and 10 (+) shows an 86bp band corresponding to the base change creating the MspI RFLP, and DNA from patient 10 also gives a 412bp band arising from the base change creating the apoB_{3500} mutation.

3.7 Implications of these findings.

The aim of this experiment was to identify common sequence differences in the region of the apoB gene encoding the primary receptor binding domain that are in linkage disequilibrium with the XbaI RFLP. This was done by comparing sequence differences between the X- alleles and the X+ alleles. It was expected that any sequence difference would not have a large impact on the serum cholesterol levels of any one individual, but since it would be common, it would be present in at least one of the 18 apoB alleles whose sequence was investigated in this study. All the patients recruited to this study demonstrated reduced binding of their LDL to a normal LDL-receptor in one or other of the assays used (Demant et al. 1988, Houlston et al. 1988, Frostegard et al. 1990). However, in none of the alleles examined were sequence differences detected in the region of the apoB gene coding for amino acids 3130-3630. It was therefore concluded that any sequence variations causing defective binding in these patients are outside this region of the apoB gene.

Other workers have demonstrated similar results in swine, where an allele of apoB, termed Lpb5, is associated with reduced clearance of LDL, hyperlipidaemia and atherogenesis. These workers sequenced the region of the gene coding for amino acids 3133 - 3497 (the primary receptor binding domain in swine apoB) and found no unique differences between this allele and three other alleles which were not associated with the same phenotype (Maeda et al. 1988, Purtell et al. 1993). This implies that sequences outside this region must also affect interaction of LDL with the LDL-receptor in pigs.

Individuals with a diagnosis of FH were excluded from this study, and none of the patients had diabetes or thyroid disfunction. Thus it is also unlikely that in these patients the

reduced binding of the LDL is due to environmental factors or gene defects other than those in the apoB gene, but at the present time there is no direct evidence for a genetic involvement in causing this reduced binding. In both the LDL-turnover studies and the U937 assays used to select these patients, care was taken to isolate LDL of the density fraction 1.025 - 1.050kg/l, thus excluding significant amounts of Lp(a) particles and lipoproteins with a high triglyceride content, since these are known to have a different affinity for the normal LDL-receptor *in vitro* (Kleinman et al. 1985).

The original receptor binding domain (amino acids 3147 - 3381) was proposed on the basis of a number of different criteria (Yang et al. 1986, Knott et al. 1986) but there are several reasons to believe that amino acid substitutions outside this originally-proposed region can affect binding of LDL to the LDL-receptor. The strongest evidence comes from the mutation at amino acid 3500. The glutamine for arginine substitution that creates this may directly affect the interaction between the ligand domain of apoB and the LDL-receptor or it may have an indirect affect on binding by altering the conformation of the protein in this region, however it has been demonstrated that this substitution virtually abolishes the binding of apoB from the mutant allele to the LDL receptor (Section 1.4.3).

Evidence that the carboxyl-terminal end of the protein is involved in modulating binding to the LDL-receptor has come from the study of a patients with premature stop codons, causing C-terminal truncated apoB protein (Section 1.3.2). One such patient has an apoB protein that is truncated at residue 4041, and is thus only 89% of the full length protein [designated B89] (Talmud et al. 1989) This patient has hypobetalipoproteinaemia due to the rapid clearance of his apoB-89 containing LDL which has increased affinity to the LDL-receptor on normal fibroblasts (Krul et al. 1989) suggesting that residues in the deleted carboxyl-terminal portion of apoB may also have a role in the interaction of apoB with its receptor. Further evidence has

come from the work of Chapman and his colleagues using monoclonal antibodies raised to short peptides of apoB, which demonstrated that an antibody raised to a peptide consisting of amino acids 4007 to 4019 will inhibit binding and degradation of LDL by normal LDL-receptors (Chapman et al. 1989). Another study (Milne et al. 1989) found that in addition to the epitopes between 3100 and 3600, an epitope between amino acids 4100 and 4300 also became unavailable to monoclonal antibodies after LDL had been bound to its receptor, indicating that this epitope is either directly involved in interaction with the receptor or that it is affected by stearic hinderance when LDL is bound. Finally, Fantappie et al. (1992) demonstrated that LDL had increased binding to the receptor after a monoclonal antibody which recognises an epitope at amino acids 4082-4306 had been bound to LDL, perhaps indicating that masking these residues may remove their inhibitory effect.

Thus, it seemed appropriate to continue the search for mutations that may affect the binding of apoB to the LDL receptor in exon 29 of the apoB gene which encodes these carboxyl-terminal residues.

CHAPTER 4

TWO AMINO ACID SUBSTITUTIONS IN APOB ARE IN COMPLETE ALLELIC ASSOCIATION WITH THE AG(X/Y) POLYMORPHISM.

Summary

An A to G base change has been identified in exon 29 of the apoB gene, this results in the substitution of serine for asparagine at residue 4311 of mature apoB-100. Huang et al. (1990) reported a C to T base change in exon 26 that causes the substitution of leucine for proline at residue 2712 of apoB. Complete allelic association was found between the alleles at both these sites and the Ag(x/y) immunochemical polymorphism of LDL in a sample of 118 Finnish individuals. This implies that either one, or both of these substitutions combined, could be the molecular basis of the Ag(x/y) antigenic determinants - the allele encoding $serine_{4311}$ plus $leucine_{2712}$ representing the Ag(x) epitope, and that encoding $asparagine_{4311}$ plus $proline_{2712}$ the Ag(y). In a sample of 86 healthy, Swedish individuals the Leu_{2712}/Ser_{4311} allele is associated with reduced serum levels of LDL cholesterol and apoB and with significantly raised levels of HDL. However these differences are of smaller effect than those associated with the XbaI RFLP of the apoB gene in this sample. 664 individuals from European, Asian, Chinese and Afro-Caribbean populations have also been genotyped and complete association has been found between the sites encoding residues 2712 and 4311 in all of these samples, although there are large differences in the allele frequencies among these populations. In addition, there is strong linkage disequilibrium with allelic association between the alleles of these sites and those of the XbaI RFLP in all the populations examined.

4.1 Identification of an A to G base change generating the Asn_{4311} to Ser substitution.

Using patients 1 to 9, described in sections 2.1.1 and 3.2, chemical cleavage mismatch analysis was used to screen amplified DNA at bases 12326 to 13379 in exon 29 of the apoB gene. This gene region encodes the epitope for the monoclonal antibody B6B3. It has been suggested that this antibody

detects a protein polymorphism (M. Tikkanen. personal communication) and so this region was chosen as the first in exon 29 to be examined. As in Chapter 3, the aim of the experiment was to search for mutations in linkage disequilibrium with the XbaI RFLP that could be the cause of differences in serum apoB and LDL cholesterol levels seen to be associated with this RFLP.

The exon 29 gene region examined was covered in two overlapping PCR fragments: **a/b** and **c/d.** The positions and sequences of the primers used are shown in Figure 4.1.1. Both PCR fragments covered the site of the EcoRI RFLP - a G to A change at base 12669, this gave the possibility of using this site as a positive control in the mismatch analysis of both PCR products. As in chapter 3, amplified, labelled DNA from patient 9 was used as the "probe" in the mismatch analysis. No cleavage products were detected in the region amplified with primers **a/b** in these 9 individuals (results not shown). All of these individuals, including the "probe" were homozygous for the EcoRI R+ allele and so it was not possible to detect a cleavage product resulting from this polymorphism. The results of screening the second amplified region: **c/d** are shown in Figure 4.1.2. In addition to the 800bp PCR product a 230bp cleavage product was seen in all the individuals examined. This was detected with hydroxylamine and corresponds to a mismatched C base in the "probe" DNA. Although present in all the individuals, it is at reduced intensity in patient 8 (the probe) and in patient 10, indicating that this base change is polymorphic and that these two individuals may be heterozygous for the change. Since the 230bp cleavage product was only detected in PCR product **c/d,** and not in **a/b,** the mismatched C detected must be 230bp from the 3' end of PCR **c/d.** Direct sequencing was thus directed at the 3' end of this PCR product. Typical results are shown in Figure 4.1.3. A G to A base change can be seen at nucleotide 13141. This will generate an Asn to Ser change at residue 4311.

Patients 8 and 9 are both heterozygous for this change, whilst patients 2 to 7 are homozygous for A at nucleotide 13441, explaining the intense cleavage band in patients 2 to 7 when probed with the heterozygous DNA from patient 9. This polymorphism was independently discovered by Navajas et al. (1990).

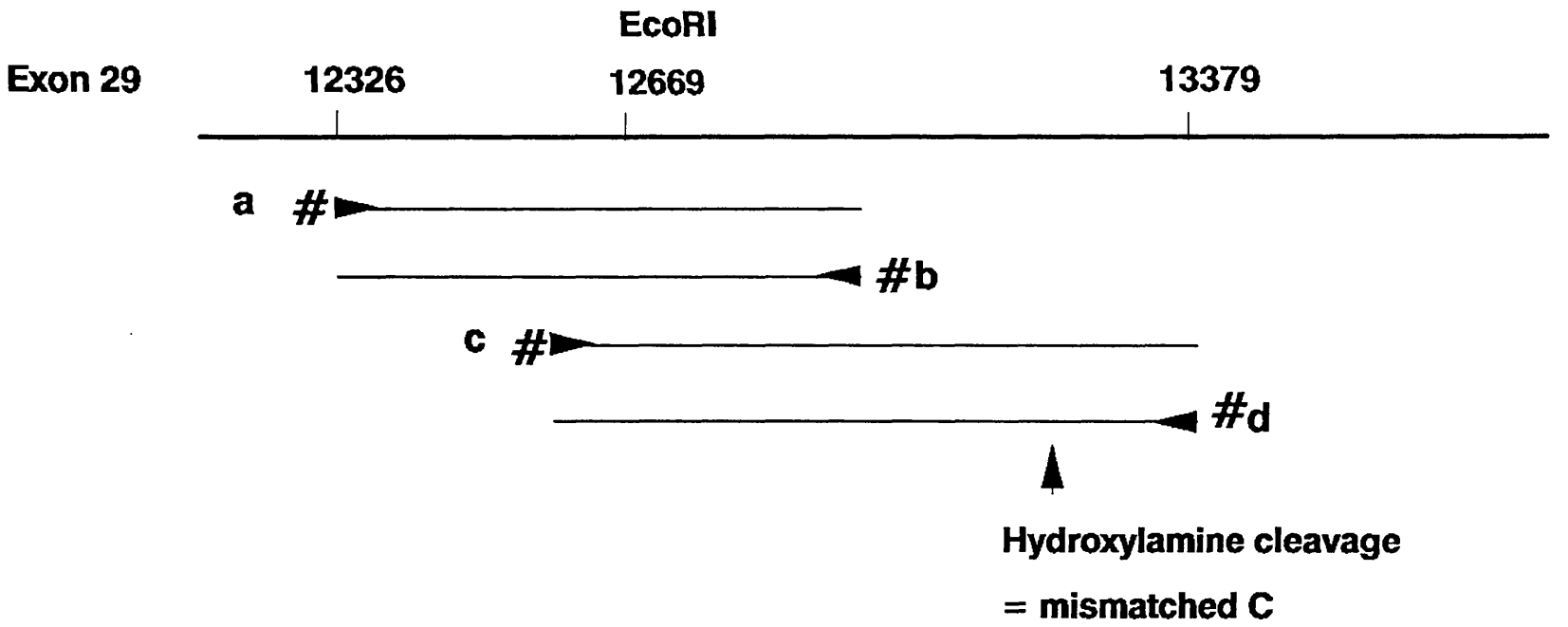


Figure 4.1.1 Partial map showing the gene region within exon 29 covered.

Nucleotide numbers are given along the top. Overlapping PCR products used in the analysis are depicted below. Positions of primers are represented by arrows. Positions of 5' end-labelling with ^{32}P are denoted by #. Primer sequences are : (5'-3') a - GCTAACCTCTCTGAAAGACAACG, b - GAAGATTACGTAGCACC TCTGTG, c -CAGGAACTGTTGACTCAGGAAGGCCAAGCC, d - AGCAACTAACAGGTTCTTGATCAGACTGAC.

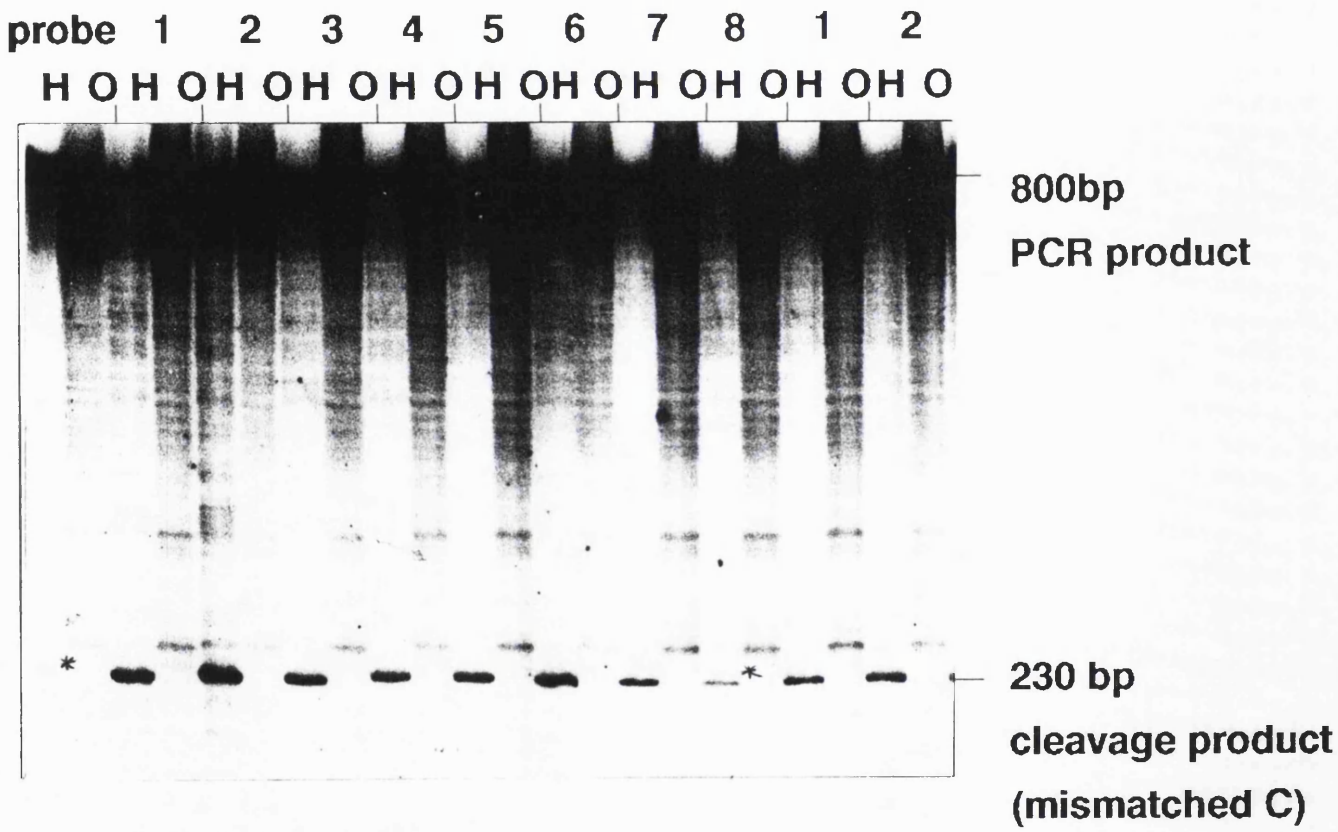


Figure 4.1.2 Chemical cleavage mismatch analysis.

Autoradiograph showing the results of the analysis. DNA samples 1 to 8 and probe [the "probe" DNA sample hybridised against itself as a negative control] are presented (samples 1 and 2 are shown twice). In each case the hydroxylamine reaction track (H), is on the left, and the osmium tetroxide reaction (O) is on the right. The 800bp uncleaved PCR product is visible along the top with the 230bp hydroxylamine cleavage product beneath. The samples yielding reduced intensity bands [ie. heterozygous for the base change] are marked *.

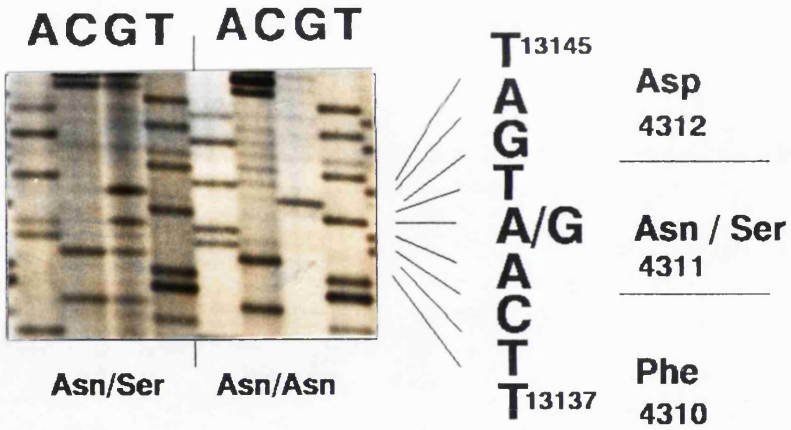


Figure 4.1.3 Direct sequencing of DNA amplified across the mismatch.

Nucleotides 13137 to 13145 (encoding amino acids 4310 to 4312) are shown for two DNA samples. The right hand four tracks show sequenced DNA from an individual homozygous for A, encoding Asn. The left hand four tracks are from an individual heterozygous for A and G , encoding Asn and Ser.

4.2 Asn_{4311} to Ser and Pro_{2712} to Leu are in complete linkage disequilibrium with Ag(x/y) in 119 Finnish individuals.

The report of a further polymorphism in apoB, this time of a C to T change generating a Pro_{2712} to Leu substitution was published by Huang et al.(1990) concurrently with the findings described in section 4.1. An investigation was thus made into whether either of these polymorphisms was associated with the Ag(x/y) protein polymorphism, since at the time of study, this was the only remaining Ag polymorphism without a candidate molecular basis. The 119 Finnish individuals used in this study are the combined North Karelia and Finnish samples (sections 2.2.3 and 2.2.4) who have all been phenotyped for all the Ag antigenic determinants using antisera (Xu et al. 1989b).

The base changes creating the two polymorphisms were detected using labelled allele-specific oligonucleotides (ASOs), following PCR amplification of the test individuals' DNA. The positions and sequences of the primers used, together with an example of the results obtained are shown in Fig 4.2.1. Detailed methodology is given in Section 2.9.1 The genotype distributions and relative allele frequencies for all the population samples studied are shown in Table 4.2.2

Genotypes at both sites were scored for each of the 119 individuals and the results compared to the Ag protein phenotyping results. Complete association was found between these two polymorphisms and Ag(x/y). Pro_{2712} is in complete allelic association with Asn_{4311} and Ag(y) and conversely, Leu_{2712} is with Ser_{4311} and Ag(x). The results are shown in Table 4.2.3. In this Finnish sample the frequency of the common allele [$Ag(y)/Pro_{2712}/Asn_{4311}$] is 0.79 and the distribution of the genotypes is as expected for a sample in Hardy-Weinberg equilibrium.

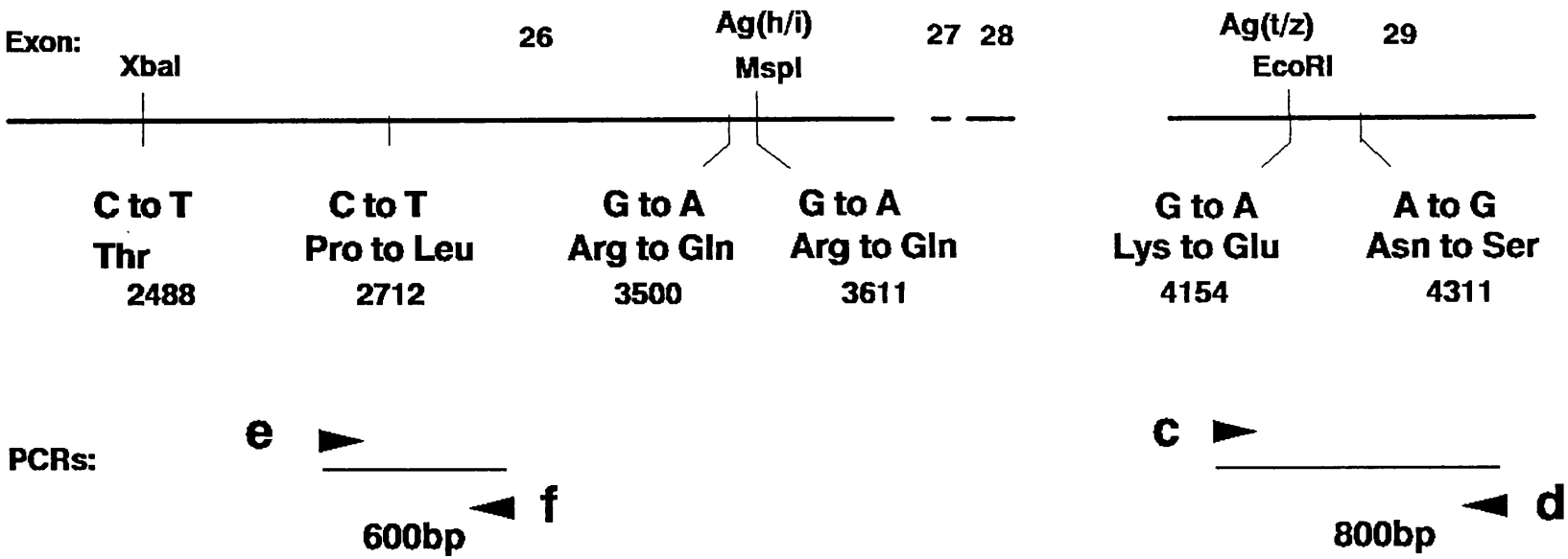
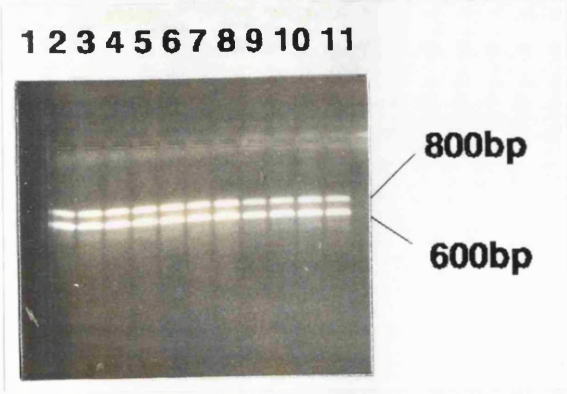


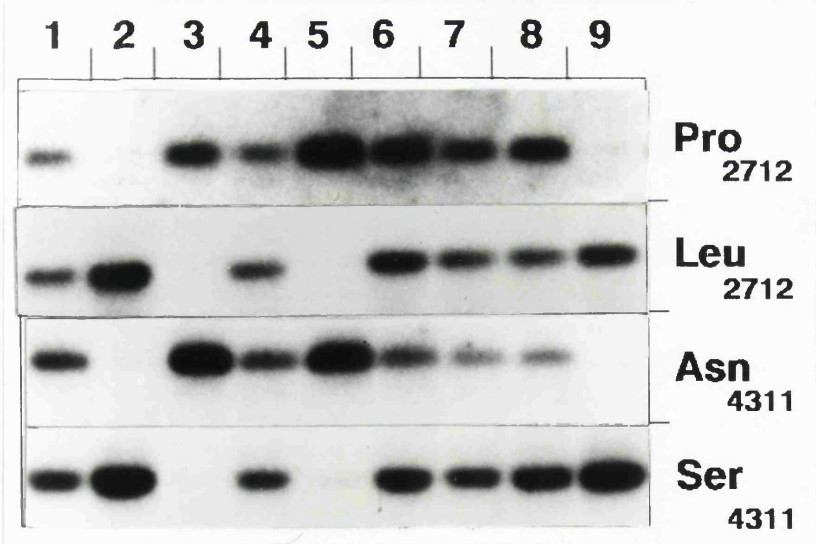
Figure 4.2.1 Allele specific oligonucleotide melting to detect Pro_{2712} to Leu and Asn_{4311} to Ser.

A Map of exons 26 to 29 showing positions of single base changes and the amino acid substitutions that they create. Where the changes also generate RFLPs and [Ag] polymorphisms these are shown along the top. The positions of the PCRs used to detect the base changes are shown below. The sequences of primers **c** and **d** are given in Figure 4.1.1. The sequences of the others are: (5' to 3') **e** - ATCATCA GAACCATTGACCAGATGCTGAAAC, **f** - TGACAATCACTCCATTACTAAGCTCCAGTG. The sequences of the ASOs are: (5' to 3') **Pro** - CACATACCAGAATTC, **Leu** - GAA TTCTAGTATGTG, **Asn** - AATCTTCAATGATTATA, and **Ser** - TATAATCACTGAAGATT.

Figure 4.2.1 continued



B Photograph of the results of the multiplex PCR using primers **c/d** (800bp fragment) and **e/f** (600bp fragment).



C Autoradiograph showing an example of the results obtained. Quadruplicate blots of amplified DNA from eleven individuals are shown. These are hybridised to the ASOs recognising the alleles for (from top to bottom): Pro_{2712}, Leu_{2712}, Asn_{4311}, Ser_{4311}. Sample 1 is an illustration of DNA from an individual heterozygous at both sites; sample 2 from one homozygous for Leu_{2712}/Ser_{4311}, and sample 3 from one homozygous for Pro_{2712}/Asn_{4311}.

Table 4.2.2 Genotype Distribution and relative allele frequencies in the population samples

	Sample size	Pro_{2712}→Leu			Asn_{4311}→Ser			Frequency of Pro_{2712}/Asn_{4311} (95%CIs)
		PP	PL	LL	NN	NS	SS	
Finnish	118	65	44	9	65	44	9	0.74 (0.66-0.82)
British	37	21	13	3	21	13	3	0.74 (0.60-0.88)
Swedish Controls	86	58	23	5	58	23	5	0.81 (0.73-0.89)
Swedish Patients	142	95	43	4	95	43	4	0.82 (0.76-0.88)
Asians	152	44	70	38	44	70	38	0.52 (0.44-0.60)
Chinese	82	7	34	41	7	34	41	0.29 (0.19-0.39)
Afro-Caribbeans	47	37	10	0	37	10	0	0.89 (0.80-0.98)

Table 4.2.3 Complete Association between Pro_{2712}→Leu, Asn_{4311}→Ser and Ag(x/y) in 119 Finnish Individuals.

	Pro_{2712} to Leu			Asn_{4311} to Ser		
Ag(x/y)▾	ProPro	ProLeu	LeuLeu	AsnAsn	AsnSer	SerSer
xx			9			9
xy		44			44	
yy	66			66		

This finding of complete allelic association indicates that either one or both of these loci may form the epitopes detected as Ag(x/y). Wu et al. (1991) independently demonstrated complete association of Pro_{2712} with Ag(y) and Leu_{2712} with Ag(x) in a sample of 18 individuals. While this finding helps to confirm mine, these authors claim that the Pro2712 to Leu substitution is responsible for the Ag(x/y) polymorphism. This statement is clearly unwarranted because of the high degree of linkage disequilibrium in this gene region. Presently it is only possible to conclude that the molecular basis of Ag(x/y) could be: the Pro_{2712} to Leu or the Asn_{4311} to Ser substitution alone, or it could be due to both sites acting together within the tertiary structure of the protein to form one epitope. It is also possible that, with this amount of linkage disequilibrium, neither of these two substitutions are the true molecular basis, but that it is due to other as yet undiscovered amino acid substitutions, in complete linkage disequilibrium with these.

4.3 Investigation of this allelic association in other human ethnic groups and in higher apes

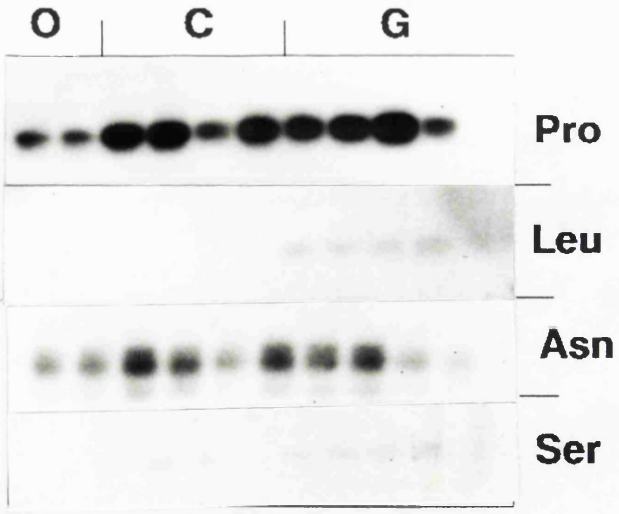
Since there was complete allelic association between these polymorphisms in the Finnish population sample it was decided to further investigate this association in other ethnic groups. Samples from Swedish, British, Asian, Chinese and Afro-caribbean populations (Sections 2.2.5 - 2.2.9) were studied using the same

ASO detection technique. Complete association was found between the alleles encoding residues 2712 and 4311 in all the individuals investigated, totalling 523 (Table 4.3.1). However the allele frequency of the Pro_{2712}/Asn_{4311} allele varied from 0.89 in the Afro-caribbean population to 0.24 in the Chinese (Table 4.2.2). This estimation of these allele frequencies is in agreement with those observed for the Ag(x/y) polymorphism in similar ethnic groups (Breguet et al. 1990, Table 5.1.1).

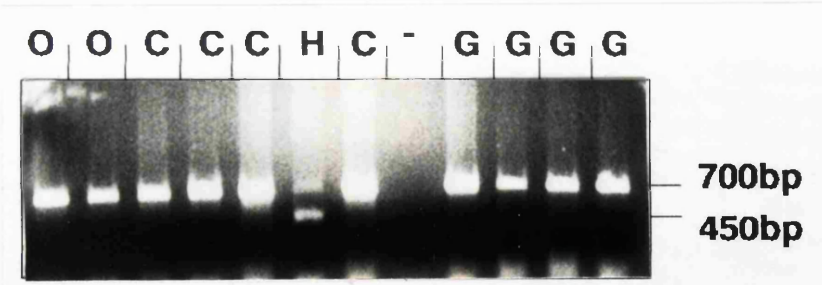
Table 4.3.1 Complete allelic association between the alleles encoding residues 2712 and 4311 in all the individuals investigated.

2712 ▼	4311 ▸	Ser/Ser	Ser/Asn	Asn/Asn	Total
Leu/Leu		100			
Pro/ Leu			237		
Pro/Pro				327	
					664

Rapacz et al.(1991) reported that in 20 chimpanzees and 6 gorillas investigated by them, the Ag(x/y) polymorphism was not apparent, all the animals were homozygous for the Ag(y) epitope. This indicates that this epitope may not be polymorphic in these species and has become so since the divergence of humans from other higher apes. Using the identical system that was developed for the human gene, the DNA from four chimpanzees, four gorillas and two orang-utans (Section 2.3) was genotyped (Figure 4.3.2). Each animal was homozygous for Pro and Asn at the apoB residues corresponding to human residues 2712 and 4311 respectively. This is compatible with these animals being homozygous for the Ag(y) epitope. At the same time these apes were genotyped for the XbaI RFLP and were all found to be homozygous for the absence of the XbaI site [the X- allele]. This indicates that this site also is not polymorphic in higher apes.



A



B

Figure 4.3.2 Genotyping the apes for Pro_{2712} to Leu, Asn_{4311} to Ser and the XbaI RFLP.

Results of typing all ten apes [2 orang-utans (O), 4 chimpanzees (C) and 4 gorillas (G)] for Pro_{2712} to Leu, Asn_{4311} to Ser are shown on the left (details as for Figure 4.2.1C) [A] and for the XbaI RFLP on the right [B]. Track H in [B] is human DNA from an individual heterozygous for the XbaI RFLP giving band of 700 and 450bp none of the ape DNA digests with XbaI (700bp band only) indicating that they are all homozygous X-X-.

Breguet et al. (1990) found Ag(x/y) to be polymorphic in all ethnic groups except native American indians. The authors thus conclude that this locus must be one of the oldest Ag variant sites, and that it probably arose before the postulated migration of *Homo sapiens* out of Africa. If this is the case then our data suggest that there has been no recombination between the sites encoding residues 2712, 4311 and Ag(x/y) in the history of modern man. I also sought, but failed to find, an intermediate combination of residues eg Pro_{2712}/Ser_{4311} or Leu_{2712}/Asn_{4311} on the same haplotype, which would have indicated the order in which the two polymorphic sites arose. One explanation for this lack of an intermediate haplotype would be if the base change creating these two sites occurred simultaneously and the two loci have not since been separated by recombination. Alternatively, the intermediate haplotype could have been lost from the gene pool by chance or due to selection. The large differences in allele frequency seen between the different ethnic groups could be explained by random genetic drift. However, it is not possible to say whether the lack of recombination is due to chance or natural selection. The lack of recombination also means that, even looking at these large and diverse samples, it cannot be determined whether the Ag(x/y) epitopes are created by residues 4311 or 2712 alone, or together, or by completely different residues. Although finding an individual with an intermediate haplotype would be helpful in distinguishing these possibilities, only *in vitro* mutagenesis and expression experiments will be able to prove the exact nature of the amino acid substitution creating Ag(x/y).

4.4 Secondary structure protein modelling

This work was done in collaboration with Drs. Gerald Laxer, Maryvonne Rosseneu and Robert Brasseur.

In order to try to elucidate what, if any, the functional effect of the Pro_{2712} to Leu, and Asn_{4311} to Ser substitutions may be, an attempt was made to determine the secondary structures of the polymorphisms of apoB using the amino acid sequence data of the relevant portions of the protein. This was done using the output of the "Predict" secondary structure prediction program and hydrophobicity profiles. This form of modelling is inherently speculative but may give some indications as to which of the amino acid substitutions, could be the molecular basis of the Ag(x/y) antigenic determinants.

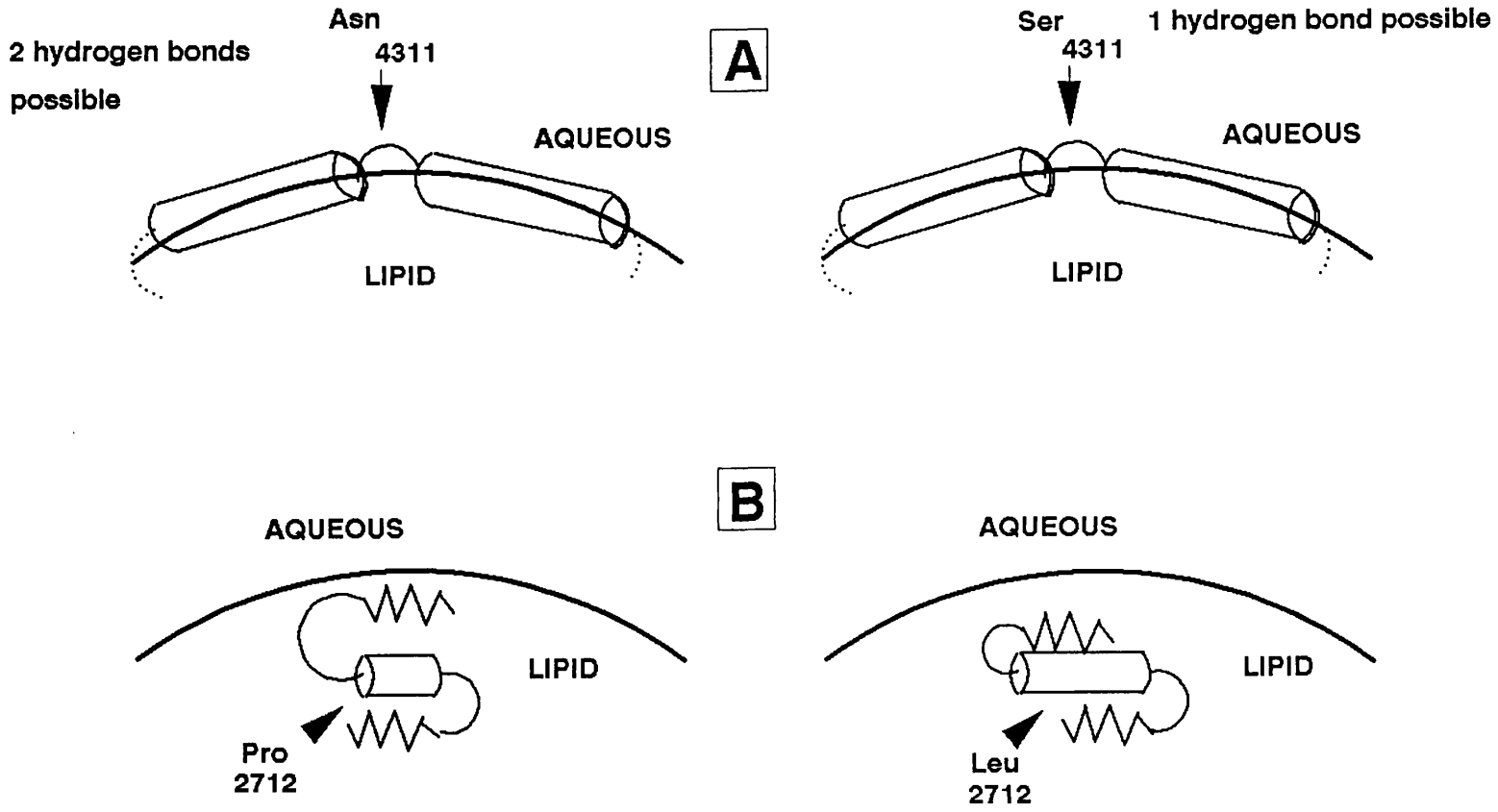


Figure 4.4.1 Cartoon representations of possible structures in the regions of residues 2712 and 4311.

Phospholipid membrane of the LDL particle is represented by the bold arcs. Potential helical regions are denoted by cylinders and likely β-sheet regions by zigzags. (Cartoons are not drawn to scale.) A Residue 4311 probably lies in a short hydrophi lic region between two longer amphipathic helices. The Asn to Ser change does not appear to alter the secondary structure, however Asn could partake in the formation of two hydrogen bonds and Ser in only one during the formation of any tertiary structure. B Residue 2712 is in a hydrophobic region generally consisting of β-sheet, however residue 2712 lies within a short helix which appears to be broken by the presence of Pro at this site but is sustained by Leu.

Residue 2712 appears to be in a very short helical section within a hydrophobic region of the protein generally comprising α-sheet and β-turn (Figure 4.4.1), and it is predicted that the presence of proline, as opposed to leucine, at this site would disturb the helical structure, but the hydrophobicity of the region would not be significantly altered by the substitution. Pro_{2712} is within one of the hydrophobic "proline-rich" consensus sequences postulated to be able to penetrate the lipid of the LDL particle (DeLoof et al. 1987a) and is thus unlikely to be exposed on the surface of the LDL particle. DeLoof and colleagues (1987a) place Pro_{2712} as the ninth residue of a 25 residue consensus motif that occurs eight times within apoB-100.

[PheGlnValProAspLeuHisIle**Pro**GluPheGlnLeuProHisIleSerHisThrIleGluValProThrPhe].

It is notable that in six of these eight motifs proline is conserved as the ninth residue, in the other two it is replaced once by asparagine and once also by leucine - which may indicate that the Pro to Leu change may be conservative within such a lipid binding region. Alternatively, this may simply reflect the hypermutability of the second C in the proline codon: C**C**N is mutated to C**T**N to generate the Lys substitution..

Residue 4311 occurs in a short joining region between two postulated, long, amphipathic helices (Figure 4.4.1), indicating that it is likely to be exposed on the surface of the LDL particle. However, the Asn_{4311} to Ser substitution is also not predicted to have a major effect on apoB conformation, unless this residue is involved in hydrogen bonding, in which case Asn could participate in the formation of two bonds and Ser in only one. Amphipathic helices are thought to have "surface-seeking" properties within the LDL particle (DeLoof et al. 1987b). This indicates that residue 4311 may be exposed on the surface of the particle and would therefore be available to act as an epitope for an antibody. Another possibility is that residues 2712 and 4311 are closely related in space upon the surface of the lipoprotein particle due to the folding of the protein. If this is so, they could jointly form the Ag(x/y) epitopes, Asn_{4311} interacting with Pro_{2712} or Ser_{4311} with Leu_{2712} as the only two functional combinations of the protein. Although it is

not yet possible to use computer modelling to explore the possibility that residues 2712 and 4311 may interact in the tertiary structure, such interaction cannot be excluded.

4.5 The Ag(x/y) epitopes are in linkage disequilibrium with the XbaI RFLP

In order to determine whether the polymorphisms at residues 2712 and 4311 are contributing to the differences in lipid and lipoprotein levels seen in association with the XbaI RFLP of apo B, the genetic relationships between these three loci were investigated.

The Finnish, Swedish, British, Asian and Chinese samples (Sections2.2.1-7 and 2.2.9) had all been previously genotyped at the XbaI site, and so these samples, totalling 562 individuals, were used in this study. The results are shown in Table 4.5.1. There is strong linkage disequilibrium between these sites ($\Delta = 0.57$, $p < 0.001$). Moreover, all the unambiguously determined Ag(x)/Leu_{2712}/Ser_{4311} alleles are associated with the X- allele. The Ag(y)/Pro_{2712}/Asn_{4311} alleles are found unequivocally on both the X+ and X- alleles. Berg et al. (1986) also found strong linkage disequilibrium between the Ag(x) and XbaI X- alleles in a sample of 56 Norwegian individuals (section 1.5.2), but these authors did not find complete allelic association, (one individual in the Berg study unequivocally carried the Ag(x) allele on an XbaI X+ allele) and the reason for this difference remains unclear. Both studies indicate that the mutations creating the Ag(x)/Leu_{2712}/Ser_{4311} allele is likely to have arisen on a chromosome carrying the X- allele, and the data given here supports the theory that they have not since been separated by recombination. Thus, the original aim of searching for common potentially functional variants of apo B, in linkage disequilibrium with the alleles of the XbaI RFLP, was successful.

Table 4.5.1 Linkage disequilibrium with Allelic Association between XbaI RFLP, Pro_{2712}→Leu and Asn_{4311}→Ser in 562 individuals from different populations.

Ag(x/y)	Xbal		
	X-X-	X-X+	X+X+
SerSer/LeuLeu	69		
SerAsn/LeuPro	99	106	
AsnAsn/ProPro	44	129	115

Δ = 0.56, 95% Confidence Intervals =0.50 & 0.61, $p < 0.001$

4.6 Association of Ag(x/y) with differences in serum lipid and lipoprotein levels.

Berg et al. (1976) reported that the Ag(x/y) polymorphism is associated differences in serum cholesterol and triglyceride levels - individuals homozygous for the Ag(y) epitope had on average 4% higher serum cholesterol and 13% higher triglyceride levels than homozygotes or heterozygotes for the Ag(x) epitope (Section 1.5.3). To confirm this association the samples of Swedish patients and controls (Sections 2.2.5 and 2.2.6) were investigated for the association of the Pro_{2712} to Leu and Asn_{4311} to Ser genotypes with serum lipid and lipoprotein levels. These associations were compared with differences in the same parameters associated with the XbaI RFLP. The amount of phenotypic variance explained by the different polymorphisms is shown in Table 4.6.1.

Table 4.6.1 Percentage of variance (R^2x100) explained by the XbaI and Pro_{2712} to Leu / Asn_{4311} to Ser genotypes in 86 Swedish control and 85 Swedish patient Individuals.

	XbaI		Pro2712 to Leu Asn4311 to Ser	
	Controls	Patients	Controls	Patients
Total Cholesterol	2.4	0.1	1.3	0.3
Total Triglyceride	1.8	0.3	0.5	1.2
LDL Cholesterol	5.0*	0.2	0.8	0.7
HDL Cholesterol	8.1**	0.5	7.7**	1.0
ApoB	6.6**	2.9	3.6	0.8

* $p < 0.1$, ** $P < 0.05$.

In the control group the XbaI RFLP accounts for 5.0% of the phenotypic variance in serum LDL cholesterol levels ($p < 0.1$) and 6.6% of the variance in serum apo B levels, ($p < 0.05$). By contrast, the Pro_{2712} to Leu/ Asn_{4311} to Ser polymorphisms explain only 0.8% of the variance seen in LDL cholesterol and 3.6% of that seen in apo B levels. The largest effect the polymorphisms is on HDL cholesterol levels, 8.1% of which are explained by the XbaI RFLP and 7.7% by the Pro_{2712} to Leu/ Asn_{4311} to Ser polymorphisms. In the patient group, these polymorphisms explain very little of the variance in any of the lipid and lipoprotein traits measured and none reach statistical significance. This result is in agreement with many of the studies which have examined XbaI genotype differences in CAD patient groups (Section 1.5.3) and is probably due to the fact that patient groups include a high proportion of people with raised lipid and apoB levels due to a number of different causes; dietary, genetic and as a result of secondary hyperlipidaemia. Thus, the small effect of variation in the apoB gene on lipid levels tends to be masked by these other factors.

When the mean levels of the lipid traits are examined by genotype in the control group (Table 4.6.2) there is a trend for both the X- allele and the Leu_{2712}/Ser_{4311}/Ag(x) allele to be associated with lower mean levels of serum LDL cholesterol and apoB and also with raised HDL levels. The XbaI X-homozygotes have 7.5% lower cholesterol, 10% lower triglyceride, 13% lower LDL-cholesterol and 11% lower apoB levels than X+X+ homozygotes. Although there is a difference in mean HDL-cholesterol levels by XbaI genotype class, it is the heterozygotes who have significantly lower levels than either homozygous class. The Leu_{2712}/Ser_{4311} homozygotes have 7% lower LDL cholesterol and 15% lower apoB levels than Pro/Asn homozygotes and in addition have 39% higher HDL cholesterol levels. Thus both the Ag(x) and X- alleles, which are in allelic association, are associated with the same trend in apoB and LDL-cholesterol levels. This result is in concordance with the many reports on the XbaI RFLP and with Berg's report on Ag(x/y) (1976) (Section 1.5.3). No differences in mean levels of lipids or lipoproteins by genotype class were seen in the patient group (results not shown).

Moreel and colleagues (1992), have also studied the effect of the Asn_{4311} to Ser polymorphism in the control groups of their larger ECTIM study - numbering 481 individuals. They find the same association of individuals homozygous for Ser_{4311} allele having lower mean levels of total cholesterol (5%), LDL-cholesterol (7%), apoB (11%) and triglycerides (20%) than individuals with one or more Asn_{4311} alleles. Thus, this larger study reveals very similar differences in apoB and LDL-cholesterol levels to my study and additionally finds differences in total cholesterol and triglyceride levels. No difference was found in HDL levels in the ECTIM study, although an increased number of particles carrying apoAI are found associated with the Ser_{4311} allele. The Moreel et al. study additionally found a significant association of the Ser_{4311} allele with a reduced number of lipoproteins carrying both apoE and apoB (LpE:B) - Ser/Ser homozygotes have 35% lower levels of these particles than individuals carrying one or more Asn allele. The findings of the ECTIM study thus confirm and extend the lipid associations that I report, they also confirm that the effect of the Ser_{4311} allele is small and apparently recessive in nature.

Table 4.6.2 Comparison of mean (+SD) serum lipid and apoB levels by genotype for the XbaI RFLP and the Pro_{2712} to Leu / Asn_{4311} to Ser polymorphisms in the Swedish Control group.

	Xbal			Pro2712 to Leu Asn4311 to Ser		
	X-X-	X-X+	X+X+	LLSS	LPSN	PPNN
Number of Individuals	18	40	25	4	23	56
Total Cholesterol mmol/l	5.79 ±1.19	5.98 ±1.20	6.26 ±1.22	6.31 ±1.2	5.84 ±1.14	6.08 ±1.20
Total Triglyceride mmol/l	1.46 ±0.70	1.44 ±0.61	1.63 ±2.25	1.58 ±0.77	1.30 ±0.51	1.58 ±1.60
LDL Cholesterol mmol/l	3.56 ±1.25	3.87 ±1.29	4.11* ±1.24	3.66 ±1.27	3.79 ±1.17	3.93 ±1.28
HDL Cholesterol mmol/l	1.50 ±1.27	1.29# ±1.23	1.44 ±1.21	1.81# ±1.29	1.35 ±1.28	1.30 ±1.21
ApoB mmol/l	99.5 ±13.5	109.7 ±20.9	112.3# 16.1	93.7 ±11.6	106.6 ±14.8	109.8 ±19.8

* $p < 0.1$, # $p < 0.05$.

The association of the Ser_{4311} allele with reduced LpE:B might indicate a mechanism through which the reduced lipid and lipoprotein levels are mediated. Particles which carry both apoE and apoB, but not apoCIII are principally IDL, thus it is likely that there are reduced numbers of these lipoproteins in Ser_{4311} homozygotes. Reduced serum numbers of these particles could be due to lower secretion rates of VLDL or by the liver, or they could indicate increased direct clearance of VLDL and IDL from the plasma, rather than the alternative hydrolysis of VLDL to IDL and finally LDL that is mediated by LPL.

It is possible to speculate that Leu_{2712}/Ser_{4311} VLDL and IDL may have increased direct clearance due to these particles either carrying increased apoE or having the apoB receptor binding domain in an accessible conformation. Both possibilities would enhance receptor-mediated clearance of such lipoproteins.

4.7 Implications of these findings.

The discovery of the $Leu_{2712}/Ser_{4311}/Ag(x)$ allele which is in allelic association with the XbaI X-allele means that the original aim of searching for common, potentially functional variants of the apoB gene, in linkage disequilibrium with the XbaI RFLP has been successful. The results of this study and those of Moreel et al (1992) demonstrate that $Leu_{2712}/Ser_{4311}/Ag(x)$ explain some, but probably not all, of the differences in serum LDL cholesterol and apoB seen in association with the XbaI RFLP. If not all the differences are explained by this allele, this suggests that there are other, as yet undetected, functional variants of apoB that must account for the remainder of the differences seen to be associated with the XbaI RFLP.

This study has also demonstrated linkage disequilibrium between two polymorphic sites: Pro_{2712} to Leu and Asn_{4311} to Ser that has apparently been maintained over a distance of 7kb of DNA through the history of modern man - approximately 100-200 thousand years. The absence of individuals with only one of these changes raises the hypothesis that the base changes creating these two polymorphic sites arose simultaneously, and that they have not since been separated by recombination. This poses the question of whether this lack of recombination is due to chance or selection. In order for selection to be the cause it would require that the recombinant alleles be dominant and deleterious. If this is the case, it is not likely that natural selection would act through altered susceptibility to coronary artery disease, only the case of homozygous FH due to LDL receptor defects could be said to reduce fitness in this manner. However in this study and others extensive searches have been made for recombinant alleles and none have so far been confirmed.

Such strong linkage disequilibrium between sites also highlights the fact that association studies alone cannot confirm whether sequence changes detected in this way are having functional effects. From

this study it is not possible to determine which of the two amino acid substitutions, if either, is the molecular basis of the Ag(x/y) epitopes. Nor is it possible to say whether they are responsible for the differences in lipid and lipoprotein levels associated with them. Only *in vitro* expression studies of the two sites, individually and in combination, can demonstrate whether these substitutions are the true causes of the differences reported.

CHAPTER 5.

A POSTULATED PHYLOGENETIC TREE FOR THE HUMAN APOB GENE.

Summary

Using published data on polymorphic sites in the apoB gene it was possible to postulate a phylogenetic tree for the 3' end of the gene, covering the time since the divergence of human beings from other primates. This simple model assumed no obligatory recombination events or multiple occurrences of the same mutation. The model was then extended to cover the entire gene and tested in the Swedish patient and control individuals. All the simple haplotypes postulated in the model were observed unequivocally, but three complex, unpredicted haplotypes were unambiguously observed and a further nine, much rarer haplotypes were deduced to occur in these samples. The frequencies of the haplotypes postulated in the model do not differ between the patient and control samples, however most of the complex haplotypes occur more frequently in the patient group than in the controls. Two of these complex haplotypes, defined by the combination of the Ag(a) epitope and the presence of the XbaI cutting site, were associated with raised serum apoB levels in the control group and significantly elevated levels in the patient group. It is proposed that these observations explain in part the consistent association reported between the XbaI polymorphic site in the apoB gene and levels of plasma lipids.

5.1 Introduction.

In Chapter 4 the unexpectedly high degree of linkage disequilibrium and allelic association between the XbaI RFLP and Ag(x/y) / Pro_{2712} to Leu / Asn_{4311} to Ser was discussed. It was also shown that the Ag(x/y) polymorphism explains some, but not all, of the lipid and lipoprotein

differences associated with the XbaI RFLP. The haplotype carrying both the XbaI X- and the Ag(x) encoding alleles is associated with reduced levels of LDL-cholesterol and apoB. This suggests that one or more other haplotypes carrying the X+ allele may be associated with a lipid raising trait and that together these haplotypes would account for all the differences associated with the XbaI RFLP.

In order to test this it is necessary to define the apoB haplotypes that exist in the general population. Breguet et al. (1990) defined haplotypes of apoB by determining the Ag antigenic determinants at the protein level, in 1464 individuals from 13 population samples representing all ethnic groups. The very high degree of linkage disequilibrium is clearly demonstrated by the fact that the five di-allelic markers, which have the potential to generate 32 haplotypes , have only been reported in 14 different haplotypes. The 14 haplotypes identified in the study are shown in Table 5.1.1. together with the number of populations in which it was observed, as a measure of its overall frequency.

Later, Rapacz et al. (1991) examined the Ag polymorphisms in chimpanzees and gorillas using a combination of protein phenotyping and DNA sequencing. The ape apoB was not polymorphic for any of the Ag antigenic determinants. The antisera recognised the epitopes Ag(g,d, y,t) but did not react with the Ag(h/i) pair. However, when the ape DNA was sequenced across the gene region encoding human residue 3611, the sequence was CGG encoding Arg = Ag(i) and so the ape haplotype was deduced to be Ag(g,d,y,i,t). This is identical to a human haplotype defined by Breguet et al. (1990), and so this haplotype (shown in Table 5.1.1) is assumed to be the ancestral haplotype from which all the other human haplotypes have arisen by a combination of mutation and recombination. Since this haplotype is most frequent among the Bantu tribe and declines in frequency in order from Africa, through Europe and Asia to Australasia, the authors state that this genocline supports an out-of-Africa migration for the human

population as opposed to the alternative multi-regional evolution theory (Stringer and Andrews 1988). This study also indicates that the Ag system of polymorphisms is unique to *Homo sapiens* and has evolved since our divergence from the common ape ancestor. Using a combination of our own data, together with that of Breguet et al. (1990) and Rapacz et al. (1991) a phylogenetic tree for the human apoB gene can be postulated which is rooted through the ancestral haplotype.

Table 5.1.1 The Ag haplotypes observed by Breguet et al. (1990)

Haplotype	Observed in no. of populations (13 investigated)	Equivalent observed in current study*
Ag(xacti)	6	-
Ag(xagti)	12	Yes 3
Ag(xagzi)	2	-
Ag(xdcti)	8	Yes 12/13
Ag(xdgti)	10 - all low freq.	-
Ag(yaczi)	1	Yes 10
Ag(yagth)	1	N/A
Ag(yagti)	12	Yes 2
Ag(yagzi)	10	Yes 4
Ag(ydcth)	9	N/A
Ag(ydcti)	13	Yes 6
Ag(ydgth)	1	N/A
Ag(ydgti) - **Ancestral**	13	Yes 1
Ag(ydgzi)	2	-

* The number given refers to the haplotype assigned in Table 5.3.4. N/A = not applicable since the Ag(h/i) epitopes were not differentiated in the current study.

5.2 Postulating the model phylogenetic tree.

Initially, only the 3' end of the gene (exons 26-29) was considered (Dunning et al. 1992) since there is clearly little recombination in this portion of the gene, making the tree relatively simple to draw. The postulated tree was the most parsimonious possible given the existing data: Commencing with the ancestral haplotype, all the mutations generating the variant sites were assumed to have arisen on a single haplotype in a sequential order. Six variant sites were considered: Pro_{2712} to Leu, Asn_{4311} to Ser [Ag(x/y)], Glu_{4154} to Lys [Ag(t/z)], Arg_{3500} to Gln [FDB-100], Arg_{3611} to Gln [Ag(h/i)] and the XbaI RFLP which does not alter residue Thr_{2488}.

The postulated tree is shown in Figure 5.2.1. The ancestral haplotype for this tree is essentially that of Rapacz et al. (1991) with the addition of data from Section 4.3 which showed that the four gorillas, four chimpanzees and two orang-utans examined are homozygous for the alleles encoding Pro_{2712}, Asn_{4311} and the XbaI X- allele, thus indicating that these alleles are present on the human ancestral haplotype.

The two haplotypes, designated A_0 and B_0 (Figure 5.2.1) differ by only the C to T base change that creates the XbaI polymorphic site. It is likely that B_0 is derived from A_0 since only A_0 is seen in the higher apes that have been examined, however both are present in major human ethnic groups although their frequencies vary widely between populations. A possible explanation for this is that the C to T change arose early in human evolution (before the migration out-of-Africa) and the frequency differences between ethnic groups are due to subsequent genetic drift. However it is also possible that the XbaI polymorphism was present in the ape ancestral populations at a low frequency. Only 38 zoo-bred, higher apes have been investigated by Rapacz et al. (1991) and in Section 4.3 and it is possible that low-frequency polymorphism has not been detected.

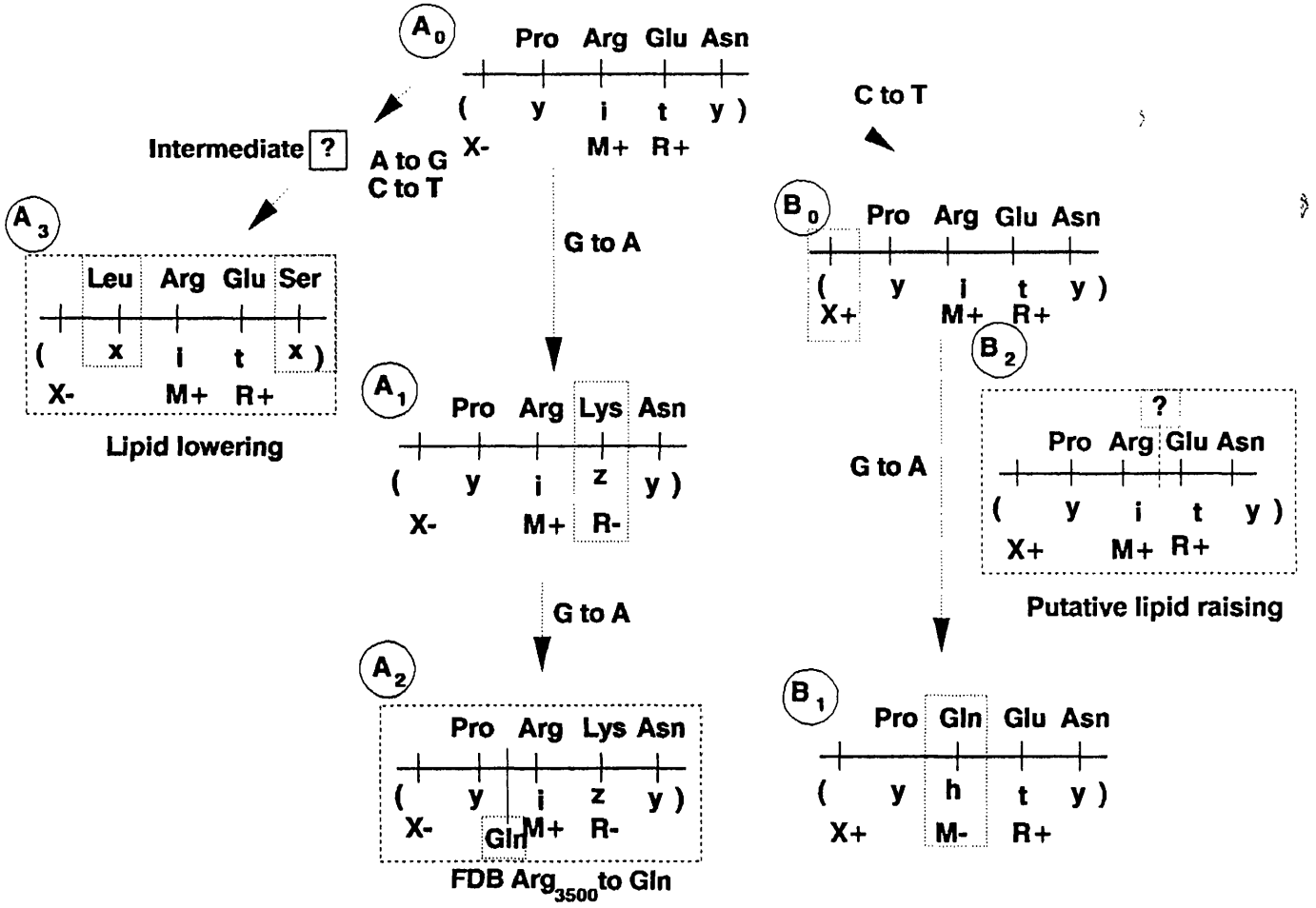


Figure 5.2.1 The postulated parsimonious tree for the 3' end of the apoB gene.

Line representations of the apoB haplotypes for the 3' end of the gene. Amino acids (three letter codes) at residues 2712,4154 and 4311 respectively are shown above the line. The presence (+) or absence (-) of the restriction enzyme sites (X= XbaI; M = MspI and E = EcoRI) and partial Ag haplotypes are shown beneath. The ancestral haplotype is shown (Ao) and all sequential haplotype have been given a letter and number code. Haplotype A_2 additionally shows the Gln residue at position 3500 that is the cause of FDB. Base changes creating the new haplotypes are shown next to the arrows. Full discussion is given in the text. The differences associated with each new haplotype are marked in boxes.

It can be inferred that all the other mutations arose on one or other of these chromosomes in a sequential order, and little or no subsequent recombination is required to explain the observed haplotypes. Haplotype A_1 can be postulated to have been created by a G to A base change on A_0, giving rise to the EcoRI RFLP at the DNA level and the Ag(t/z) polymorphism at the protein level. The G to A mutation that generated Arg_{3500} to Gln (designated A_2 in Fig 5.2.1) appears to have arisen on an A_1 chromosome, since all the patients, but two, so far identified with this disease have inherited this haplotype (Section.1.4.3, Table 1.4.3).

The only exception to the pattern that each haplotype differs from the previous one by a single base change is Ag(xit) (designated A_3 on Figure 5.2.1), which differs from it progenitor A_0 at the two sites: Pro_{2712} to Leu and Asn_{4311} to Ser. Studies to-date have not revealed an intermediate haplotype (Sections 4.2 and 4.3) implying that these two base changes may have arisen simultaneously or within a short time interval. Alternatively, due to chance or selection, the intermediate haplotype may now be extremely rare in the populations sampled and therefore has not been detected.

The Ag(h/i) polymorphism, which is the MspI RFLP at the DNA level, and appears to be the result of the Arg_{3611} to Gln amino acid substitution, may have been created by a G to A base change on haplotype B_0 giving rise to B_1, since this is the most frequently observed haplotype carrying this change reported by Breguet et al. (1990) (Table 5.1.1).

To explain the data from Section 4.6 an additional, functionally distinct, haplotype must be postulated (designated B_2) that has been created by an as yet unidentified sequence change occurring on the B_0 haplotype. This sequence change would be in allelic association with the XbaI X+ site and have the effect of raising serum cholesterol levels, thus accounting for the

residual phenotypic variance explained by the XbaI RFLP but not by the Leu_{2712}/Ser_{4311}/Ag(x) allele (Sections 1.5.3 and 4.6).

Multiple occurrences of the same mutation, reverse mutation and recombination were not considered at this stage. However the data of Breguet et al. (1990) (Table 5.1.1) demonstrate that one or more of these events must have occurred, since the Ag(x), Ag(z) and Ag(h) alleles are all observed on more than one haplotype. However, in each case, one haplotype carrying these epitopes is significantly more frequent world-wide than the others and this is the haplotype presented in this tree.

5.3 Extending and testing the postulated tree.

This postulated tree was next extended to cover the gene region 5' of the XbaI site. Again a combination of our own data and that of Breguet et al. (1990) was utilised. There is some evidence for recombination having occurred at this end of the gene, but again the most frequently observed combinations of alleles were judged to form the main branches of the tree. The extended tree is shown in Figure 5.3.1. Haplotype 1 is the ancestral haplotype. A base change on this haplotype creating the Ala_{591} to Val substitution [Ag(a/d) epitope change] is postulated to have occurred at an early stage in modern human evolution to generate haplotype 2 and then all the subsequent haplotypes on this branch of the tree also encode the Ag(a) epitope. A study of the sample of Finnish individuals (Xu et al. 1989) indicated almost complete allelic association between the Ag(c) epitope and the signal peptide deletion allele (Δ), making it impossible to deduce which of these two mutations arose first. However, since both show allelic association with the XbaI X+ allele, it can be inferred that they occurred on the X+ branch of the tree to create haplotype 6.

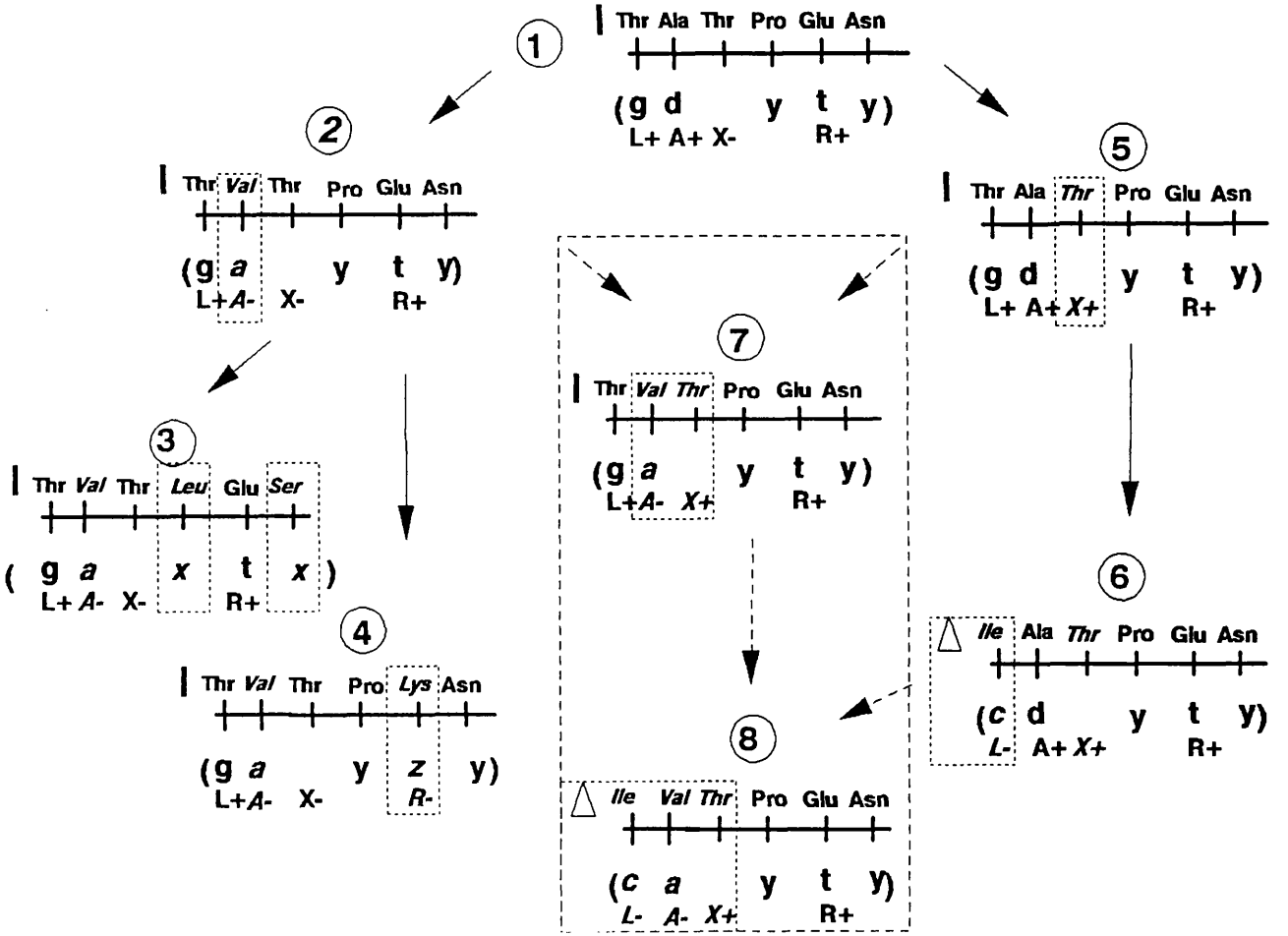


Figure 5.3.1 The extended phylogenetic tree for the apoB gene.

The most parsimonious tree is shown around the outer edge (Haplotypes 1 to 6). Haplotype 1 is the deduced ancestral haplotype. Within the box are shown the two most frequently unambiguously observed complex haplotypes (Haplotypes 7 and 8), together with arrows showing from which other haplotypes they may be derived. Amino acids (three letter codes) at residues 71, 591, 2488, 2712, 4154 and 4311 respectively are shown above the line. The signal peptide Insertion/Deletion polymorphism is represented by I and Δ. The epitopes of the Ag polymorphisms for which the amino acids are candidate molecular bases are shown below the line, and the presence (+) or absence (-) of RFLP sites below these. (L = ApaLI, A = AluI, X = XbaI, R = EcoRI). The differences creating each new haplotype are shown in boxes

The validity of the extended tree was tested using the Swedish patient and control groups (Sections 2.2.5 and 2.2.6). Seven polymorphic sites were considered in testing the tree: the signal peptide three amino acid deletion (I/Δ), Thr_{71} to Ile [Ag(c/g)], Ala_{591} to Val [Ag(a/d)], Pro_{2712} to Leu, Asn_{4311} to Ser [Ag(x/y)], Glu_{4154} to Lys [Ag(t/z)], and the XbaI RFLP. Genotyping of these samples for the signal peptide Insertion/Deletion (I/Δ) polymorphism, Pro_{2712} to Leu, Asn_{4311} to Ser [Ag(x/y)], Glu_{4154} to Lys [Ag(t/z)], and the XbaI RFLP have been described previously (Peacock et al. 1991, Section 4.2). Details of genotyping for Thr_{71} to Ile [Ag(g/c)] and Val_{591} to Ala [Ag(a/d)] are given in Section 2.9.3 and typical results are shown in Figure 5.3.2. The Arg_{3500} to Gln change was not present in the patient group (Tybjaerg-Hansen et al. 1990) and the Arg_{3611} to Gln change [Ag(h/i)] had not been determined in these samples and so these sites were not considered in testing the tree.

The complete data for all the Swedish patients and controls examined are shown in Appendix I. The genotype frequencies are given in Table 5.3.3, and they are as expected for a sample in Hardy-Weinberg equilibrium. The degrees of pair-wise linkage disequilibria between each of the polymorphisms is shown in Figure 5.3.4.

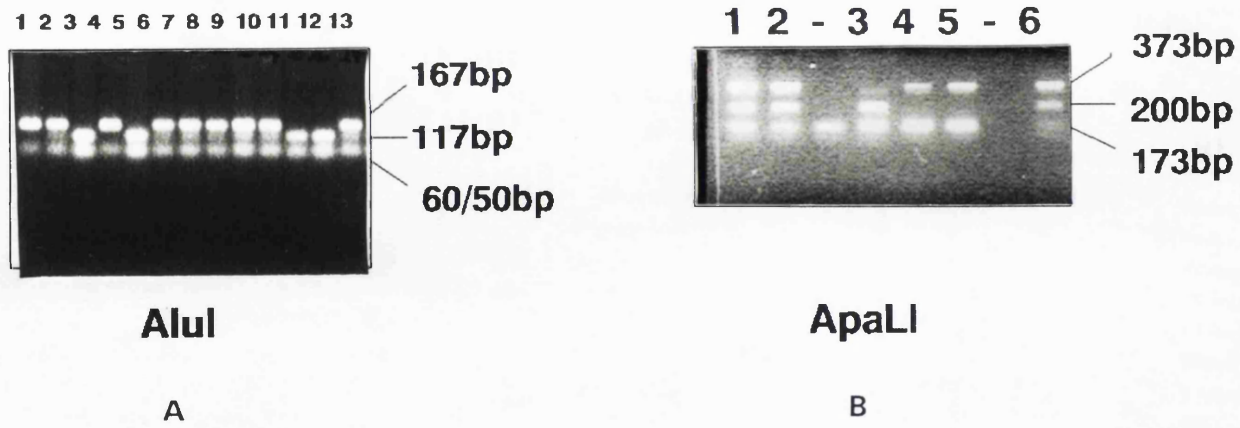


Figure 5.3.2 Examples of digestions to detect AluI and ApaLI RFLPs.

A The AluI digest detects Val_{591} to Ala [Ag(a/d)]. The 60bp fragment is constant. 167bp fragment = A- = Val_{591}, 117 + 50bp fragments = A+ = Ala_{591}. Tracks **1**, **2** and **3** show DNA from a homozygote A-, a heterozygote anda homozygote A+. B ApaLI detects the Thr_{71} to Ile [Ag(c./g)] polymorphism. 373bp band = L- allele = Ile_{71}, 200 + 173bp bands = L+ allele = Thr_{71}. Tracks **1** and **2** show DNA from heterozygotes, track **3** shows DNA from a homozygous L+ individual and tracks **4** and **5** from an L- homozygote.

Table 5.3.3 Rare allele frequencies for the polymorphisms studied in the Swedish patient and control groups

	Signal Peptide Ins/Deletion Δ	Ag(c/g) Ag(c) = Ile_{71}	Ag(a/d) Ag(a) = Val_{591}	XbaI X-	Ag(x/y) Ag(x) = Leu_{2712}/Ser_{4311}	Ag(t/z) Ag(z) = Lys_{4154}
Controls	0.31	0.30	0.55	0.45	0.18	0.18
[95% CIs]	[0.30-0.32]	[0.29-0.31]	[0.54-0.56]	[0.44-0.46]	[0.17-0.19]	[0.17-0.19]
(No of Individs)	(90)	(87)	(90)	(90)	(90)	(90)
Patients	0.38	0.37	#0.45	0.47	0.19	0.16
[95% CIs]	[0.373-0.387]	[0.363-0.377]	[0.443-0.457]	[0.463-0.477]	[0.185-0.195]	[0.155-0.165]
(No of Individs)	(144)	(142)	(142)	(143)	(141)	(143)

# $p < 0.05$

95% CI for difference = 0.007 - 0.193

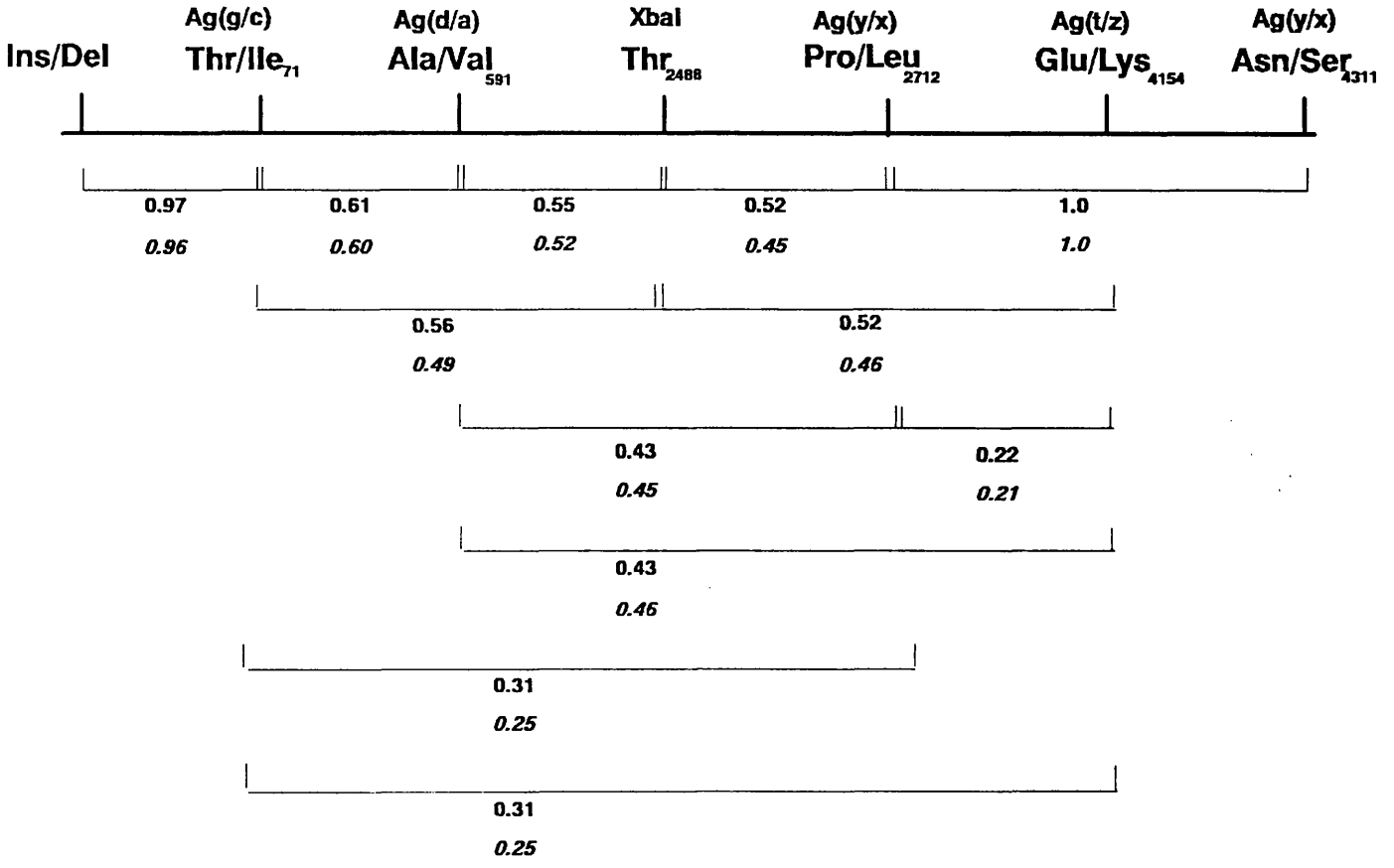


Figure 5.3.4 Pair-wise linkage disequilibria for each of the polymorphisms studied.

Correlation coefficients (Δ) are shown for all the pairs. The upper figure is that for the control group, and the lower figure (*in italics*) for the patient group. All values are significantly higher than zero ($p < 0.05$).

Table 5.3.5 Unequivocally observed and assigned haplotypes in the Swedish patient and control groups

	Controls		Patients		Total
	Unequivocal	Assigned	Unequivocal	Assigned	
1 Igd X- yt	1	8	11	11	30
2 Iga X- yt	2	3	1	6	13
3 Iga X- xt	13	19	7	38	76
4 Iga X- yz	10	23	7	35	75
5 Igd X+yt	10	12	21	11	54
6 ΔcdX+yt	19	30	48	42	135
7 Iga X+ yt	8	15	9	16	46
8 ΔcaX+yt	3	1	0	2	6
9 ΔcdX-yt	0	1	4	5	10
10 ΔcaX+yz				3	3
11 Δcd X- yz				1	1
12 Δcd X- xt				1	1
13 Δcd X+ xt				2	2
14 Δga X+ yt		1			1
15 Δgd X+ yt				1	1
16 Δga X- xt		1		2	3
17 Icd X+ yt				1	1
18 Igd X- yz				1	1
Total:	66	114	108	178	466

Differences in haplotype frequencies between control and patient groups:-

Haplotypes 7 and 8 27/180 vs 27/286, Chi-square = 3.25, $p < 0.1$.

Haplotypes 9 to 18 3/180 vs 21/286, Chi-square = 7.30, $p < 0.01$.

The genotypes of all the individuals were assigned to haplotypes, commencing with those which could be determined unambiguously *ie.* the individuals who were homozygous at either all of the polymorphic sites or all except one. The haplotypes that were assigned in this manner fell into nine groups (Table 5.3.5, haplotypes 1 to 9). Six of these nine types (haplotypes 1 to 6) are simple and were predicted to exist in the most parsimonious phylogenetic tree model - assuming no recombination events between sites. 82% of all the haplotypes characterised were classified into haplotypes 1 to 6. However, in order to explain the other three unambiguously determined haplotypes: **7** IgaX+yt, **8** ΔcaX+yt and **9** ΔcdX-yt it is necessary to propose either the occurrence of recombination events within the apoB gene, or some other mechanism such as multiple occurrences of the same mutation.

The haplotypes of the remaining individuals were assigned by fitting their genotypes firstly to the six haplotypes proposed in the most simple model and secondly to the other three unequivocally characterised haplotypes. This left 2/90 (2.2%) individuals from the control group and 14/144 (9.7%) from the patient group whose haplotypes could not be assigned to groups that had already been demonstrated to occur unambiguously. For these individuals an attempt was made to assign one of the individual's two chromosomes to a known haplotype leaving another previously unobserved one, these are shown in Table 5.3.5 (haplotypes 10-18). Although this system necessarily involves estimation and therefore inherent errors, three of the eight haplotypes deduced in this manner apparently occur more than once (haplotypes 10, 13 and 16) suggesting that they may be real. It is notable that the frequencies of these assigned haplotypes [haplotypes 9 to 18] are significantly higher in the patient group (21/286 = 0.073) than in the control group (3/180 = 0.017).

In most cases the haplotypes can be arranged so that each could have been generated by

a single base change occurring on the preceding haplotype, giving rise to a single amino acid substitution. There are two haplotypes where this is not the case: Haplotype 3 differs from its progenitor, haplotype 2, at two sites (Section 5.2), and haplotype 6 [ΔcdX+yt] differs from its proposed precursor, haplotype 5 [IgdX+yt] by the substitution of Ile for Thr at residue 71 and also by the three amino acid deletion in the signal peptide (Δ). There are some individuals with haplotypes carrying a signal peptide deletion allele in conjunction with an allele encoding Ag(g), or the insertion allele (I) with Ag(c) (Table 5.3.5) which might indicate an intermediate haplotype, the frequency of such chromosomes in the samples examined to date is low - 7/466 = 1.5% in this study, and 2/106 = 1.9% in Xu et al. (1990).

5.4 Comparison of this model with other published haplotype data.

The data on Ag haplotypes in Breguet et al. (1990) covers population samples from all human ethnic groups and was used with data from this laboratory in postulating the phylogenetic tree. Thus, the branches of the tree represent the most frequent haplotypes seen world-wide. However this model has only, as yet, been tested in the Swedish population which may not be typical of other ethnic groups. Another method of checking the validity of the unequivocal and assigned haplotypes in the Swedish population is to compare them with other published genotype and haplotype data.

From Table 5.1.1 it can be seen that seven of a potential 11 haplotypes described by Breguet and co-workers are observed in the Swedish population - the remaining three Breguet haplotypes carried Ag(h) which was not examined in this study. The equivalent haplotypes, seen in both studies are the ones I have designated 1, 2, 3, 4, 6, 10 and 12/13. The haplotype that I designated number 5 is different from its progenitor by only the XbaI X+ allele and so this is not represented in Table 5.1.1. In another study, Ludwig and McCarthy (1990) investigated apoB

haplotypes in eight American kindreds with FDB-100. Using 10 polymorphic markers on 67 chromosomes, these workers found 22 different haplotypes. The only chromosomes that these workers observed more than twice correspond to the haplotypes that I have numbered 1 to 6. They also detected two chromosomes that correspond to haplotype 7, and at least one chromosome corresponding to each of the rare, unpredicted haplotypes that I designated 9, 10, 12, 13, 16 and 18. Thus, in addition to the haplotypes that I observed unequivocally, this family study also demonstrates the existence of five of the nine haplotypes that I was only able to deduce to be present.

Table 5.1.1 also demonstrates that Ag(c/g) is polymorphic in all ethnic groups, and all except one (native American Indian) exhibit the Ag(x/y) and Ag(a/d) polymorphisms. Ag(t/z) is not polymorphic among the American Indians or the Papuans, and Ag(h) of the Ag(h/i) pair is not detected in Tibet, China, Bali or Eastern Indonesia (Breguet et al. 1990). If the migration out-of-Africa hypothesis is accepted, this data is compatible with all the Ag polymorphisms having arisen before that migration, since the modern Africans and almost all other populations have these polymorphic sites. It would appear that either the ancestors of the eastern populations did not carry Ag(h) epitope or that it has been subsequently lost from this gene pool. In addition, the Krahó native American indians appear to have lost much of their apoB polymorphism, perhaps indicating that they are derived from a very small group of ancestors who passed through an evolutionary bottle-neck.

None of the evolutionary data on apoB is incompatible with the hypothesis of an human ancestor arising in Africa and diversifying there for a considerable period of time before a significant group migrated north and then east or west to eventually populate the rest of the world. This is consistent with the data from mitochondrial DNA studies (Cann et al. 1987). It does not appear to support the alternative multi-regional evolution hypothesis (Stringer and

Andrews, 1988) which suggests that racial differences between east and west are due to the independent evolution of ancestral groups of *Homo erectus* into separate races of *Homo sapiens* with some gene flow between the races preventing speciation. This would have generated populations with greatest differences between the furthest separated groups and few genetic differences within Africa - the opposite of what is seen. All the data also support the evidence that haplotype 1 [Ag(ydgti)] is the human ancestral haplotype from which all the others are derived (Rapacz et al. 1991).

5.5 Frequency differences between the Swedish patient and control groups.

In Table 5.3.5 haplotypes 1 to 6 were defined as simple - those that could be explained as being due to a single or double mutational event having occurred on a previously existing haplotype without recombination events. The existence of haplotypes 7 to 18, on the other hand, can only be explained by events such as recombination, multiple occurrence of the same base change or reverse mutation and so these were defined as complex haplotypes. All the simple haplotypes together with haplotypes 7, 8 and 9 were observed unambiguously, this is in part a reflection of their frequency. The complex haplotypes are much rarer than the simple ones in both groups, although the method used to fit the haplotypes to the model tree may have exaggerated this effect. However haplotypes 9 to 18 are significantly more frequent in the patient group [21/286 = 0.073] than the control group [3/180 = 0.017], Chi-square = 7.30, $p < 0.01$.

When the frequencies of the individual polymorphisms are considered [Table 5.3.3] it can be seen that the Ag(a) allele [Val_{591}] is more frequent in the control group [0.55] than among the patients [0.45], $p < 0.05$. There are no significant difference in the frequencies of any of the other individual polymorphisms between patients and controls. The reason for this frequency

difference is not immediately apparent but, if confirmed in other studies, it could be of clinical importance.

5.6 Association of the haplotypes with differences in serum lipid levels.

Having defined and assigned the different haplotypes it was then possible to test to see which, if any, were associated with differences in serum lipid levels. The effects of the XbaI RFLP on serum apoB levels in these Swedish samples were investigated in Section 4.6 and the results are shown in Tables 4.6.1 and 4.6.2. In these samples the alleles of the XbaI RFLP are associated with a significant effect on apoB levels in the group of control individuals, with a similar trend on LDL cholesterol levels, the presence of the XbaI cutting site being associated with higher mean levels of both traits. There is a non-significant trend for the $Leu_{2712}/Ser_{4311}[Ag(x)]$ to be associated with lower mean levels of serum LDL-cholesterol and apoB.

Table 5.6.1 Percentage phenotypic variance of serum apoB explained by the individual polymorphisms / haplotypes in the apoB gene.

	Controls	Patients
XbaI	6.6#	2.9
Ala_{591}-Val [Ag(a/d)]	2.2	0.8
Pro_{2712}-Leu/ Asn_{4311} -Ser **Haplotype 3**	3.6	1.0
Glu_{4154}- Lys [Ag(t/z)] **Haplotype 4**	2.3	1.6
Thr_{71} -Ile [Ag(c/g)] Signal Peptide Ins/Del **Haplotype 6**	4.0	0.3
Haplotypes 7 & 8	2.5	8.2#

# $p < 0.05$ Results are shown after adjustment for age and BMI (Controls) and age, sex and BMI (Patients).

To expand these observations, the amount of phenotypic variance of serum apoB levels due to each of the individual polymorphisms was considered separately and the results are shown in Table 5.6.1. None, except the XbaI RFLP, were associated with a significant differences in apoB levels in either the patient or control groups when considered individually. Two of these alleles, the Ag(x) and Ag(z), uniquely define single haplotypes (Haplotypes 3 and 4, Figure 5.3.1) and so these two haplotypes may be discounted from being associated with differences in apoB levels in these Swedish population samples. Similarly, the signal peptide deletion allele [Δ] and the Ag(c) allele, are predominantly associated with haplotype 6 and so it can also be deduced that haplotype 6 is not associated with a significant effect on mean apoB levels.

Since three of the unequivocally observed haplotypes [7, 8 and 9] were complex and did not fit the most simple model phylogenetic tree, the effects of these haplotypes on plasma apoB levels were next investigated. The haplotype defined by the presence of Δ and Ag(c) in conjunction with the XbaI X- allele [haplotype 9] occurred too rarely (1/180 in the control group and 9/286 in the patient group) for statistical analysis. The other two haplotypes [7 and 8] which are both defined by the presence of Ag(a) in combination with the XbaI X+ allele have combined frequencies of 27/180 (0.15) in the control and 25/286 (0.09) in the patient group. The amount of phenotypic variance explained by these haplotypes is shown at the bottom of Table 5.6.1. These haplotypes account for 8.2% of the variance in the patient group ($p = 0.007$) and 2.5% in the control group ($p = 0.15$) When the mean apoB levels in the groups of individuals with and without these haplotypes were compared, haplotypes 7 and 8 were associated with 13% higher mean apoB levels in the patient group and with 6% higher levels in the control group. These results are shown in Table 5.6.2. This difference reached statistical significance in the patient group. Individuals with these haplotypes also had raised LDL cholesterol levels in both the patient and control groups but these differences did not reach statistical significance. Individuals with haplotypes 7 and 8 were compared by quartiles of apoB levels [Figure 5.6.3] and there are a higher than expected frequency of individuals with these haplotypes in the upper quartiles, 42% (9/21) of individuals with haplotypes 7 and 8 in the control group have apoB levels above the 75th percentile ($p < 0.05$), and in the patient group 44% (7/16) have apoB levels above the 75th percentile ($p < 0.1$). This supports the hypothesis that, in both patient and control groups, these haplotypes are associated with elevated plasma apoB levels. These haplotypes thus, have a mild, but consistent co-dominant effect with more than 40% of individuals carrying these haplotypes having apoB levels in the upper quartile, and 25% having levels above the 90th percentile. This implies that these haplotypes encode common variant forms of apoB protein with a mild functional impact on the individual but a significant effect at the population level.

The fact that the effect on apoB levels associated with haplotypes 7 and 8 is greater in the patient sample than in the sample of healthy individuals is unexplained, but if confirmed is of obvious clinical significance. The effect of these haplotypes is also greater than the effect of the Xbai RFLP in the patient group. In the control group haplotypes 3, 7 and 8 together account for 6.1% of the variance in apoB levels while the XbaI RFLP accounts for 6.6%. This is an indication that these three haplotypes together explain the XbaI effect, although this will need confirmation in larger studies. There is also a trend for haplotypes 7 and 8 to be associated with raised serum LDL cholesterol levels, although these do not reach statistical significance in the patients or the controls. This is consistent with the hypothesis that whilst apolipoprotein levels are largely under genetic control, lipid and lipoprotein levels are affected to a greater extent by environmental factors (Reilly et al. 1990).

Table 5.6.2 Mean apoB levels mg/dl in patients and controls grouped by XbaI genotype or presence/absence of haplotypes 7 and 8.

Xbal genotype	Controls		Patients	
	Number	Mean ±SD	Number	Mean ± SD
X-X-	18	99.5 ± 13.5	24	127.4 ± 15.9
X-X+	40	109.7 ± 20.9	43	120.6 ± 22.7
X+X+	25	112.3 ± 16.1	26	128.3 ± 23.5
	F = 2.84 **p = 0.05**		F = 1.35 p = 0.27	
Haplotypes 7 and 8	**Controls**		**Patients**	
	Number	Mean ± SD	Number	Mean ± SD
Haplotypes 7 & 8	21	113.3 ± 19.9	16	137.5 ± 26.3
Others	62	106.6 ± 18.0	72	121.3 ± 19.8
	F = 2.06 p = 0.15		F = 7.73 **p = 0.007**	

Only male Patients were considered in this analysis.

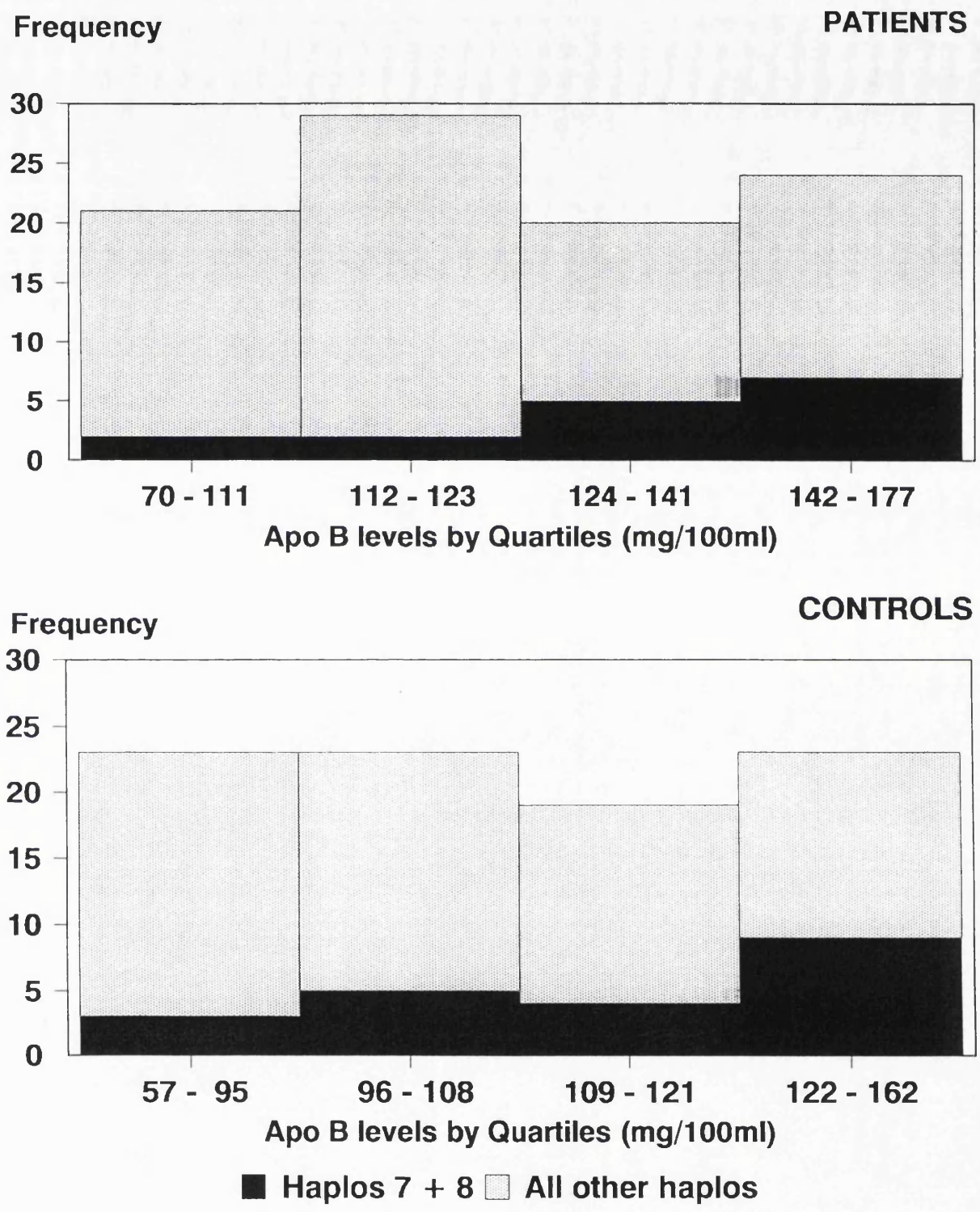


Figure 5.6.3 Analysis of ApoB Levels Within the Two Groups by Quartiles.

Black boxes represent individuals carrying one or two copies of haplotypes 7 and 8 and grey boxes represent individuals with all other haplotypes.

5.7 Implications of this study.

In this study an attempt was made to understand the mechanism of association of the XbaI RFLP with differences in serum lipid and lipoprotein levels and some haplotypes which may account for these differences have been detected. The haplotype carrying the Ag(x) and XbaI X- alleles is associated with a small mean lowering effect on LDL-cholesterol and apoB, acting in a recessive manner. By contrast haplotypes, uniquely defined by the Ag(a) and XbaI X+ alleles, are associated with mean raised levels of apoB with a mild but dominant effect. It will be of interest to see if these associations are also confirmed in other population samples that demonstrated the XbaI association, and to examine their effect at the molecular level.

During the course of this study, aspects of the evolutionary structure of the apoB gene that account for the reported strong linkage disequilibria and allelic associations have been revealed. It is the very existence of this high degree of linkage disequilibrium that has enabled haplotypes to be deduced without family studies. In general, the evidence indicates that novel mutation, rather than recombination, is the driving force for change, particularly in the 3' end of the apoB gene. This poses the question of whether this lack of recombination is due to chance or selection. In order for selection to be the cause, it would require that the forms of apoB created by recombination would be deleterious.

This study has also revealed the existence of more complex apoB haplotypes, the existence of which cannot be explained by simple novel mutation. There are several possible mechanisms by which these complex haplotypes could have occurred. Recombination events between two different apoB alleles is the most likely possibility, but it is also possible that some of the polymorphisms considered here are at hypermutable sites and that either the same mutation has

occurred twice on two different haplotypes, or that a mutation having occurred once, has subsequently reverted back to its original form. If recombination is the mechanism by which these unusual haplotypes are generated, it is notable that in all except one of the complex haplotypes seen, the site of recombination would appear to be 5' to the XbaI site. It is notable that more than 99% of the intron sequences within the apoB gene are also located 5' to the XbaI site and exon 26 (Blackhart et al. 1986) This implies that apparent recombination events may correlate with the presence of intron sequences, as suggested previously by Gilbert (1985).

The higher frequency of the complex haplotypes among the Swedish patient group, as compared to the controls, raises the possibility that recombination events within the apoB gene may be deleterious. This would be of obvious clinical importance and, if confirmed, these individuals should be investigated for dislipidaemia or other clinical risk factors. However, it is hard to imagine that increased risk of cardiovascular disease could affect Darwinian fitness, leading to a reduced overall frequency of complex haplotypes since this type of late-onset disease seldom affects people before they have reproduced. If selection is to be invoked it is perhaps more likely to have been mediated though differential lipid metabolism and resistance to famine or malnutrition during human evolution. Alternatively, there is evidence that apoB-containing lipoproteins have a role in neutralising endotoxin relaeased by gram-negative bacteria (Liaow and Floren 1992). It is possible that some variants of apoB that are better at neutralising endotoxin, enhance resistance to certain bacterial infections, which would be another mechanism through which natural selection could act. It is not inconceivable that the apoB alleles that confer resitance to famine or bacterial infection, are the same ones that now predispose to hyperlipidaemia under dietary excess.

If haplotypes 7 and 8 do have a functional effect on serum apoB levels, the question is raised as to what is the mechanism of this effect. It could be due to the presence of another

mutation, also carried on haplotypes 7 and 8, that either alone or in combination with the Ala_{591} to Val change, alters the metabolism of the LDL particle or affects the production of apoB protein. This could be one of the apoB variants listed in Table 1.5.2 or may not yet have been discovered. Alternatively, if haplotypes 7 and 8 arose by recombination between other haplotypes, the functional effect might be due to the new combinations of existing amino acid polymorphisms in the apoB protein. Evidence from a pig model for hyperlipidaemia supports this hypothesis. A group of pigs with spontaneous hyperlipidaemia and reduced FCR of LDL have been examined (Checovitch et al. 1988, Lowe et al. 1988). Although this phenotype is associated with one allele of apoB, termed Lpb5, which has been extensively sequenced, no single amino acid change is unique to this allele (Purtell et al. 1993). Instead, the hyperlipidaemia appears to be due to the unique combination of Asp_{3164} and Ala_{3447} within this allele. This hypothesis assumes that the apoB protein has a significant tertiary structure within the lipoprotein particle, for which there is growing evidence (Walsh et al. 1990, Chatterton et al. 1991, Koduri et al. 1991).

CHAPTER 6

ARG_{4243} TO THR - A NEW VARIANT IN EXON 29.

Summary

A new variant has been found in a search of exon 29 of the apoB gene. A G to C base-change at nucleotide 12937 results in the substitution of threonine for arginine at residue 4243. Protein secondary structure modelling predicts that the Thr_{4243} variant would have reduced local net positive charge and increased hydrophobicity. The carrier frequency of Thr_{4243} is estimated to be 0.035 - no homozygotes have yet been detected. There is evidence that carriers of Thr_{4243} may be more frequent among CAD patients than among controls and that carriers may have 4-20% higher serum apoB levels than Arg_{4243} homozygotes, but these differences do not reach statistical significance in the samples presented here. Twelve of the thirteen Thr_{4243} alleles examined to date share a common apoB haplotype and thus may be identical by descent.

6.1 Background

Whilst working in this laboratory, Dr. Helmut Renges sought new apoB variants in individuals with high or low binding of LDL to the LDL-receptor as measured by the U937 cell growth assay (Frostegård et al. (1990). Twelve individuals were investigated: seven with reduced receptor-binding and five with increased - binding and biochemical details for these individuals are given in Table 2.1.1. Dr. Renges used a combination of chemical cleavage mismatch analysis and (SSCP) to investigate the gene region encoding the primary receptor binding domain (using the system described in Chapter 3) and the whole of exon 29. No variants were found in exon 26 in these individuals. Only one new variant was discovered in exon 29 - an extra band in the hydroxylamine track of the mismatch analysis of PCR fragments **c/d** (Figure 4.1.1) in individual 15 (Table 2.1.1). Shown in Figure 6.1.3, this band corresponds to a mis-paired C in the "probe" DNA approximately 350bp from one end of the amplified DNA fragment.

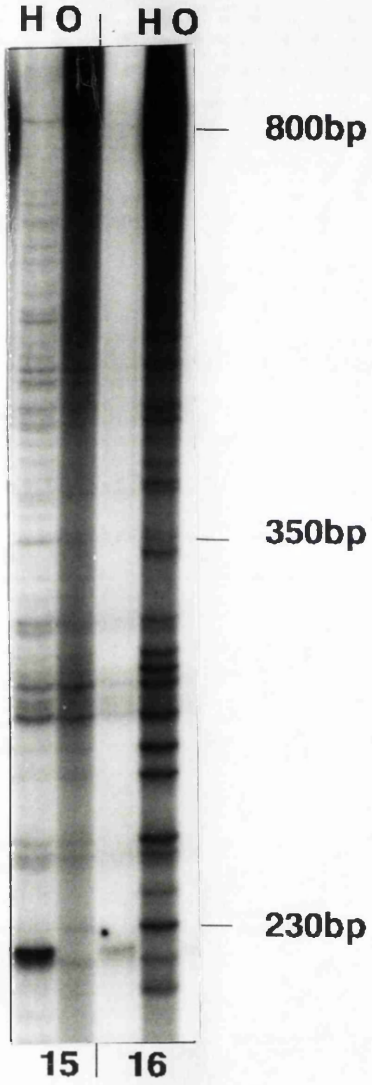


Figure 6.1.3 Chemical Cleavage Mismatch Analysis .

The DNA from individuals **15** and **16** is shown. There are two tracks per person from the hydroxylamine reaction (**H**) and the osmium tetroxide reaction (**O**). The uncleaved full-length PCR gives a band of **800bp**. The **230bp** band corresponds to the Asn_{311} to Ser change shown in Figure 4.1.2. The new **350bp** band, corresponding to a mismatched C, is in the hydroxylamine track of individual 15. ***Autoradiograph courtesy of Dr. H Renges.***

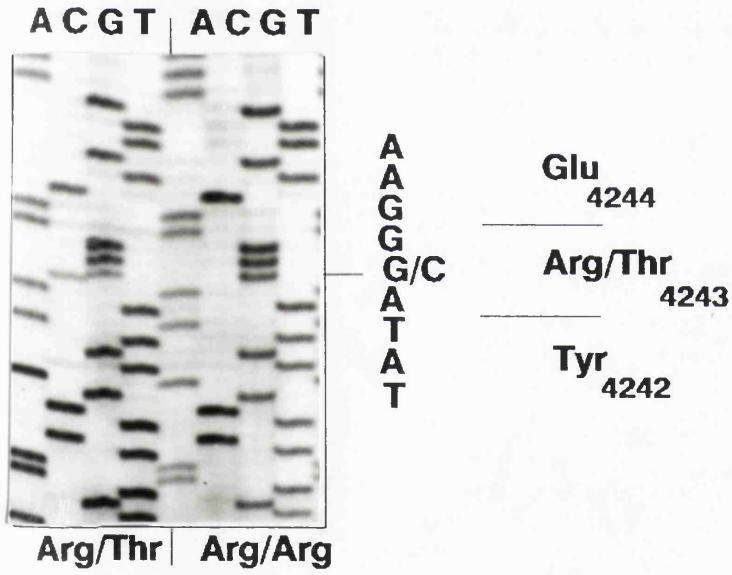


Figure 6.2.1 Direct Sequencing across the mismatch.

Nucleotides 12935 to 12941 (encoding amino acids 4242 - 4244) are shown for two DNA samples. The first four tracks show DNA from individual 15, heterozygous for G and C encoding **Arg** and **Thr**. The second set of tracks are from an individual homozygous for G encoding **Arg**.

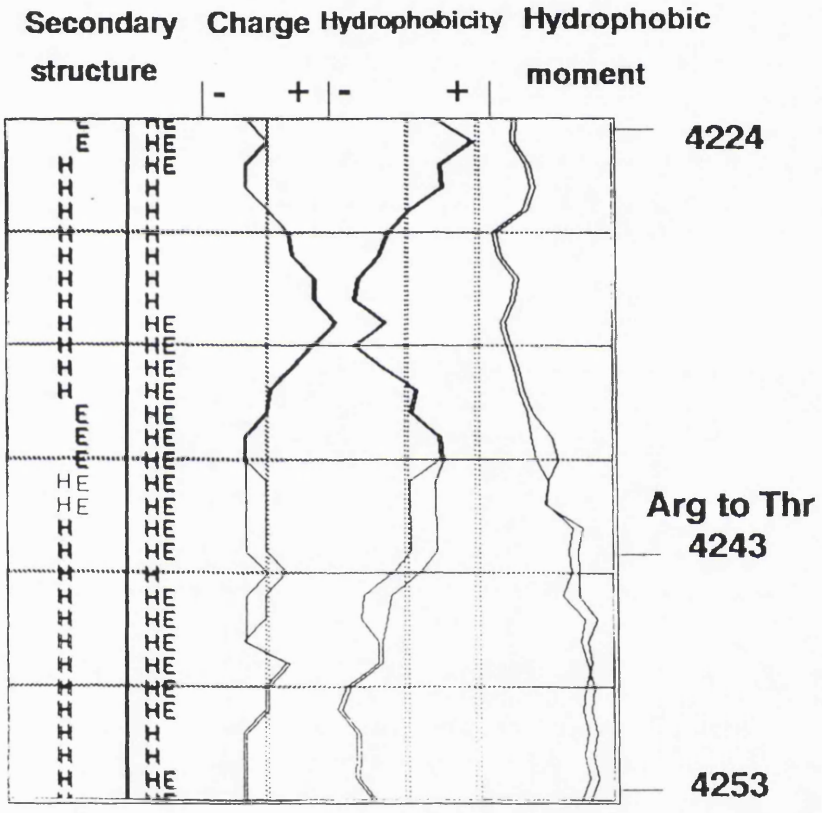


Figure 6.2.2 Protein secondary structure modelling for the Arg $_{4243}$ to Thr substitution.

The results from analysis using the DNA-Star 'Protein' program are shown for apoB residues 4224 to 4253. The **secondary structure** prediction is shown as H = helix or E = extend. Predictions of relative local **charge** (+/-), **hydrophobicity** (+/-) and **hydrophobic moment** are also shown. The Thr$_{4243}$ variant is predicted to decrease the local charge and increase the hydrophobicity.

6.2 The Sequence Change and Protein Secondary Structure Modelling.

After Dr. Renges had left the laboratory I sequenced across the region of the mismatch in individual 15 using the direct sequencing procedure described in Section 2.8.3. The results of the sequencing reaction are shown in Figure 6.2.1. The mismatch band corresponds to a G to C change at residue 12937 - this changes the codon for residue 4243 from AGG to ACG and would result in the substitution of threonine for arginine at this residue.

The effect of this amino acid substitution on local charge, hydrophobicity and helix formation in the protein secondary structure was investigated using the DNA-Star prediction program. The results are shown in Figure 6.2.2. The Arg to Thr substitution replaces a positively charged hydrophil ic residue with a neutral one in the more N-terminal of the two long amphipathic helices in this region (Figure 4.4.1). This is predicted to reduce the local charge and increase the local hydrophobicity. The predicted effect on the secondary structure would be to reduce the amount of helix and alter the pattern of amphipathicity. Although the same caution that applied to the protein modelling in section 4.4 also has to be applied here, it would appear that this is not a conservative change by conventional secondary structure modelling principles. By comparison with the known effect of such changes in apoE (Section 1.2.7), the loss of the positive charge could be predicted to have an effect. Such a change could alter the affinity of apoB for the receptor or for arterial-wall proteoglycans (Section 1.2.6), or it could affect the process of LDL oxidation (Section 1.2.4). Alternatively, the loss of the Arg residue may affect hydrogen bonding within the tertiary structure of apoB.

6.3 The frequency of Thr_{4243} in different population samples.

In order to ascertain the frequency of the Thr_{4243} variant in different population samples,

a screening system was devised using ASO melting on pools of five DNA samples amplified by PCR in the same reaction. This enabled ten alleles to be screened at a time. Pools which yielded a positive result were then expanded with each DNA sample re-amplified individually to determine which carried the variant allele.

PCR was carried out using primers **c/d** (Figure 4.1.1). After blotting, alleles were detected using oligonucleotides 5'end-labelled with fluorescein dUTP and the ECL system (Amersham UK) using the method described in Section 2.9.2. Oligonucleotide sequences and typical results are shown in Figure 6.3.1. The population samples screened were: Glasgow (Section 2.2.1), Sheffield (2.2.2), Swedish controls (2.2.5) and patients (2.2.6), STARS (2.2.10), Camberley healthy men (2.2.11), Strasbourg ECTIM patients (2.2.12) and controls (2.2.13). The frequency of the Thr_{4243} allele in the different sample is shown in Table 6.3.2. The overall allele frequency of Thr_{4243} is 0.017 - 37 heterozygotes and no homozygotes were found among 1055 individuals screened. Less than one homozygote would be expected in a population this size, for an allele frequency of 0.017, and so this genotype distribution is in Hardy-Weinberg equilibrium.

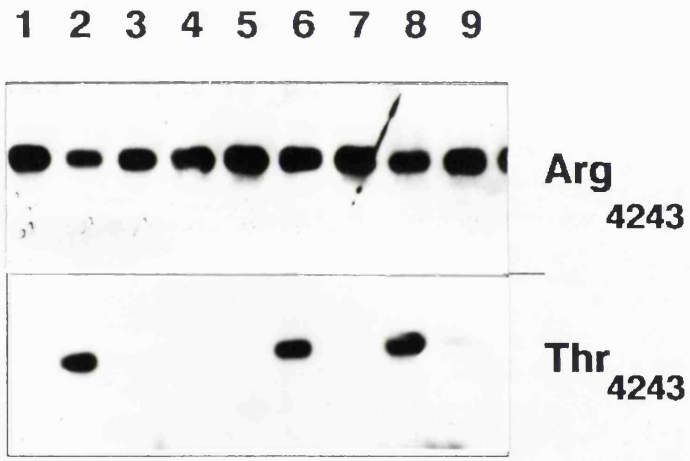


Figure 6.3.1 Allele specific oligo melting to detect Arg_{4243} to Thr.

Autoradiograph showing typical results obtained. Duplicate blots of amplified DNA from nine individuals is shown. These are hybridised to ASOs recognising the alleles for: (5' to 3') Arg_{4243} - ATGTATAGGGAACTG and Thr_{4243} - ATGTATACGGAACTG. Samples **2**, **6** and **8** are carriers of the Thr_{4243} allele, the others are homozygous for Arg_{4243}.

Table 6.3.2 The carrier frequency of Thr_{4243} in the population samples studied.

Sample	Individuals studied	Thr_{4243} carriers observed	Carrier frequency (95% CIs)
Sheffield	22	1	0.045 (0.026-0.46)
Glasgow	15	3	0.20* (0.147-0.252)
Swedish controls	93	1	0.01 (0.008-0.012)
Swedish patients	143	5	0.035 (0.032-0.038)
Camberley	366	12	0.033 (0.032-0.034)
STARS	70	3	0.043 (0.037-0.049)
ECTIM controls	170	5	0.030 (0.028-0.032)
ECTIM patients	184	7	0.038 (0.036-0.040)
TOTAL	**1070**	**37**	**0.035**

* $p < 0.001$,

Only in the Glasgow group does the carrier frequency of Thr_{4243} differ significantly from the overall frequency (0.2 vs 0.035, $p < 0.01$). The reason for this difference is unclear without examining a larger sample of hyperlipidaemic patients from Glasgow - it could be due to chance in such a small population sample, or it could indicate that this allele has a higher frequency in Scotland.

Table 6.3.3 Comparison of carrier numbers between patients and controls.

	Patients	Controls
Arg/Arg$_{4243}$	394	611
Thr/Arg$_{4243}$	18	18
Total	412	629

Chi-square = 1.73, $p > 0.1$. Allele Frequency Difference between patients and controls = 0.0075, 95% CIs for difference = (-0.004 - 0.019).

It is possible to compare the allele frequencies between patient and controls by combining all the patient groups investigated in this study: Glasgow, Swedish, STARS, and ECTIM; and all the control groups: Camberley, Swedish and ECTIM. The result of this is shown in Table 6.3.3. There is no significant difference between carrier frequencies in patients and controls ($p > 0.1$). However there is a consistent trend in that the frequency of Thr$_{4243}$ carriers is higher in the patient groups compared to the controls. Given a large enough sample population this difference may become significant, however, it is unlikely that this would be of clinical importance in most populations since the difference in frequency appears so small. However, this may be an indication that the Thr$_{4243}$ allele could interact with other genetic or environmental risk factors to generate a higher-risk phenotype.

6.4 Lipid and lipoprotein levels in individuals with the Thr$_{4243}$ allele.

Where available, lipid and lipoprotein parameters were investigated in the individuals carrying Thr$_{4243}$ to see if this allele was associated with any difference that would be clinically relevant. In these population samples there were too few carriers of this allele for one-way analysis of variance, but an attempt has made to compare the means within samples and the results are presented in Table 6.4.1. Parameters presented have not been adjusted for age or BMI.

Table 6.4.1 Comparison of mean lipid and lipoprotein levels (+ SD) between carriers of Thr_{4243} and the rest of the population sample.

Sample	Total Chol.	LDL Chol.	HDL Chol.	ApoB
Swedish controls 1	**5.77** 6.11 ±1.18	**3.54** 3.99 ±0.93	**1.47** 1.42 ±0.34	**113** 108 ± 19
Swedish patients 5	**7.03±0.79** 7.17 ±1.32	**4.81±0.87** 4.81 ±1.14	**1.12±0.31** 1.15 ±1.26	**130±7.7** 125 ±23
Sheffield 1	**4.87**	**2.22**	**2.10**	-
Glasgow 3	**8.23±0.23** 7.67 ±0.75	**6.06 ±0.20** 5.42 ±0.85	-	**193±18.5** 177 ±27.94
Camberley 12	**5.50±1.18** 5.80 ±1.82	-	-	-
ECTIM controls 5	**2.47±0.28** 2.28 ±0.46	**1.51±0.24** 1.49 ±0.38	**0.50±0.07** 0.49 ±0.15	**160±0.49** 134 ±46
ECTIM patients 6	**1.99±0.44** 2.16 ±0.62	**1.36±0.33** 1.48 ±0.48	**0.39±0.09** 0.40 ±0.13	**114±0.22** 129 ±41

Notes: No. of Thr_{4243} carriers in each sample are shown in bold in column 1. Mean for the Thr_{4243} carriers is shown in **bold** in each cell, mean for the whole sample is shown underneath.

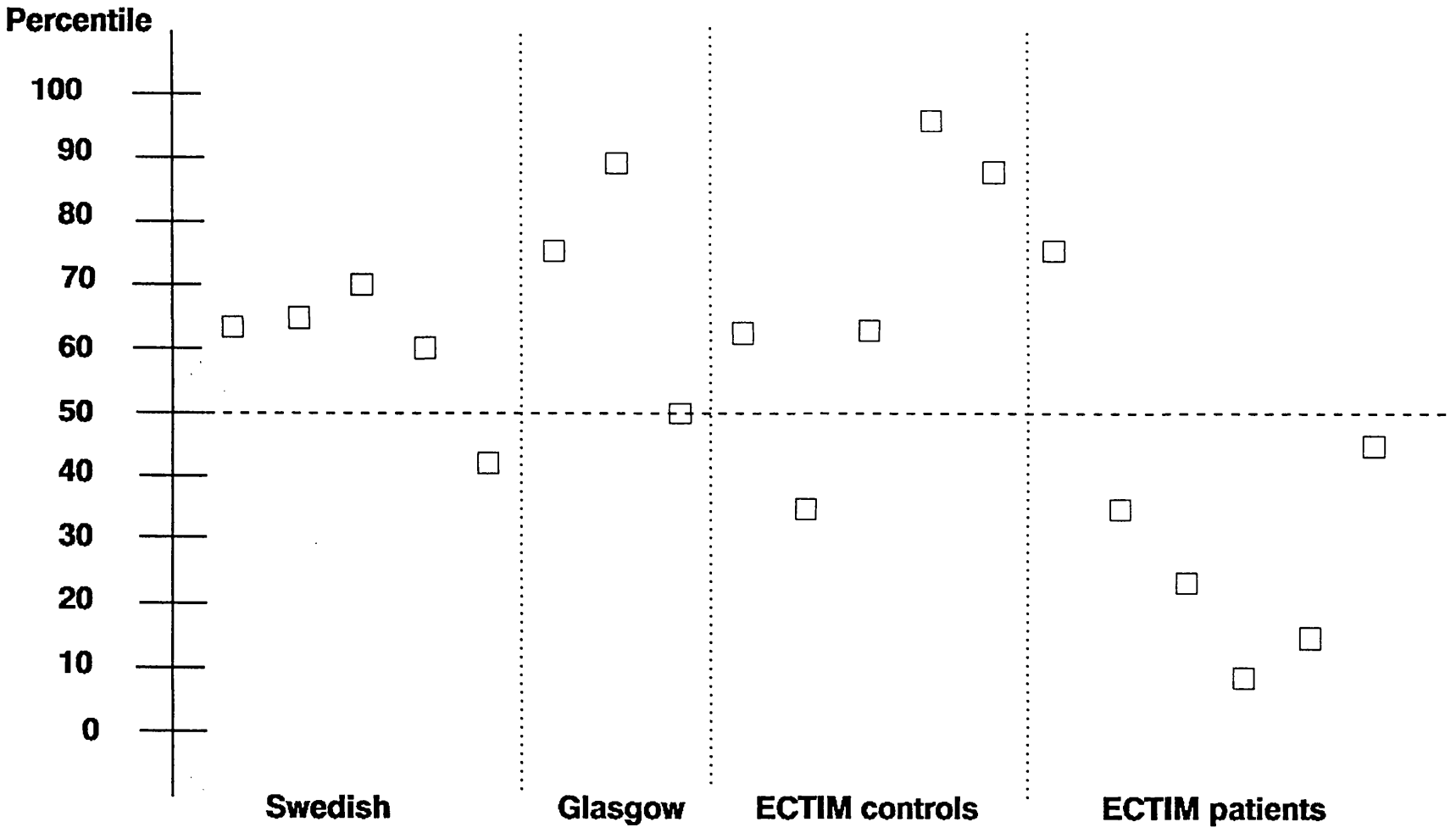


Figure 6.4.2 The serum apoB levels of Thr_{4243} carriers shown as percentiles relative to their population samples.

There is a pattern, in the samples where levels have been determined, for the Thr_{4243} carriers to have higher mean serum apoB levels than the sample mean. This trend is shown graphically in Figure 6.4.2 where the apoB levels of the carriers have been plotted as percentiles for each population sample. In the Glasgow, ECTIM control and the Swedish patient and control samples, carriers of the Thr_{4243} allele have serum apoB levels 4 - 19% higher than the population mean and 11/13 of them have apoB levels above the 50th percentile for their population sample. This effect is not, however, seen in the ECTIM patient group these results may have been affected by the fact that most of this group are receiving lipid-lowering medication. No consistent trends are seen in any of the other lipid parameters. This trend, whilst of interest, requires confirmation in large, homogeneous samples so that sufficient Thr_{4243} alleles can be detected to allow for one-way analysis of variance, enabling correction for confounding factors such as age, sex, and BMI. Alternatively a family study may enable a pattern of co-segregation of raised apoB levels with the Thr_{4243} allele to be seen. We are currently trying to obtain the families of some of these carriers in order to investigate the co-segregation of lipid, lipoprotein and LDL-receptor binding parameters with this allele.

6.5 The Thr_{4243} allele may be carried on a common haplotype.

There is sufficient genotype data on several of the individuals identified in this study to make an attempt at assigning haplotypes to the carriers of the Thr_{4243} allele. The genotypes and haplotypes of the individuals with this rare allele are given in Table 6.5.1. Twelve of the thirteen samples, in whom it is possible to assign haplotypes, share a common haplotype **1**. It is therefore likely that these 12 individuals share a common ancestor from whom they inherited this apoB variant. The remaining individual, from the Swedish patient group, carries haplotypes **9** and **16,** as characterised in Table 5.3.5. Both these haplotypes are classified as complex *ie.* they were

not predicted to exist in the most simple phylogenetic tree postulated in Section 5.3. However haplotype **9** shares markers with the 5' end of haplotype **6** and the 3' end of haplotype **1**, indicating that it may have arisen through a recombination event between these two haplotypes. Thus, if the Thr_{4243} allele were carried on haplotype **9**, this would not be inconsistent with the Thr_{4243} variant having originally arisen on haplotype **1** prior to the recombination event occurring. It is thus possible that this substitution may have only occurred once in the Western European population. Alternatively, the evidence for the existence of this variant on two different haplotypes could indicate that the mutation creating it has arisen twice independently. The Thr_{4243} variant can now be tentatively added to the phylogenetic tree. and is shown in Figure 6.5.2.

Table 6.5.1 The assigned haplotypes of the individuals carrying Thr_{4243}

Individual	Genotype							Assigned Haplo.
	Sig. Pep	71 Apa	591 AluI	2488 XbaI	2712	4154 EcoRI	4311	
SWCON 1	IΔ	TI	AA	X-X+	PP	EE	NN	1,6
SWPAT 2	II	TT	AV	X-X-	PP	EK	NN	1,4
SWPAT 3	ΔΔ	TI	AV	X-X-	PL	EE	NS	9,16
SWPAT 4	II	TT	AV	X-X-	PL	EE	NS	1,3
SWPAT 5	II	TT	AA	X-X+	PP	EE	NN	1,5
SWPAT 6	II	TT	AV	X-X+	PP	EE	NN	1,7
Sheffield 7		TT	AV	X-X	PL	EE	NS	1,3
Glasgow 8				X-X-	PL	EE	NS	1,3
Glasgow 9				X-X-	PL	EE	NS	1,3
Glasgow 10				X-X+	PP	EE	NN	1,5/6
Stars 11	IΔ		AA	X-X+				1,6
Stars 12	II		AA	X-X+				1,5
Stars 13	II		AV	X-X-				1,2

The first column shows the sample from which the individual was identified. SWCON = Swedish control, SWPAT = Swedish patient sample. Sig. Pep = signal peptide insertion(I)/deletion(Δ). Other polymorphisms are denoted by the residue affected (given along the top) and the amino acids present, in one letter code, in the body of the table. For individuals 7 -13 haplotypes were deduced from partial genotype data..

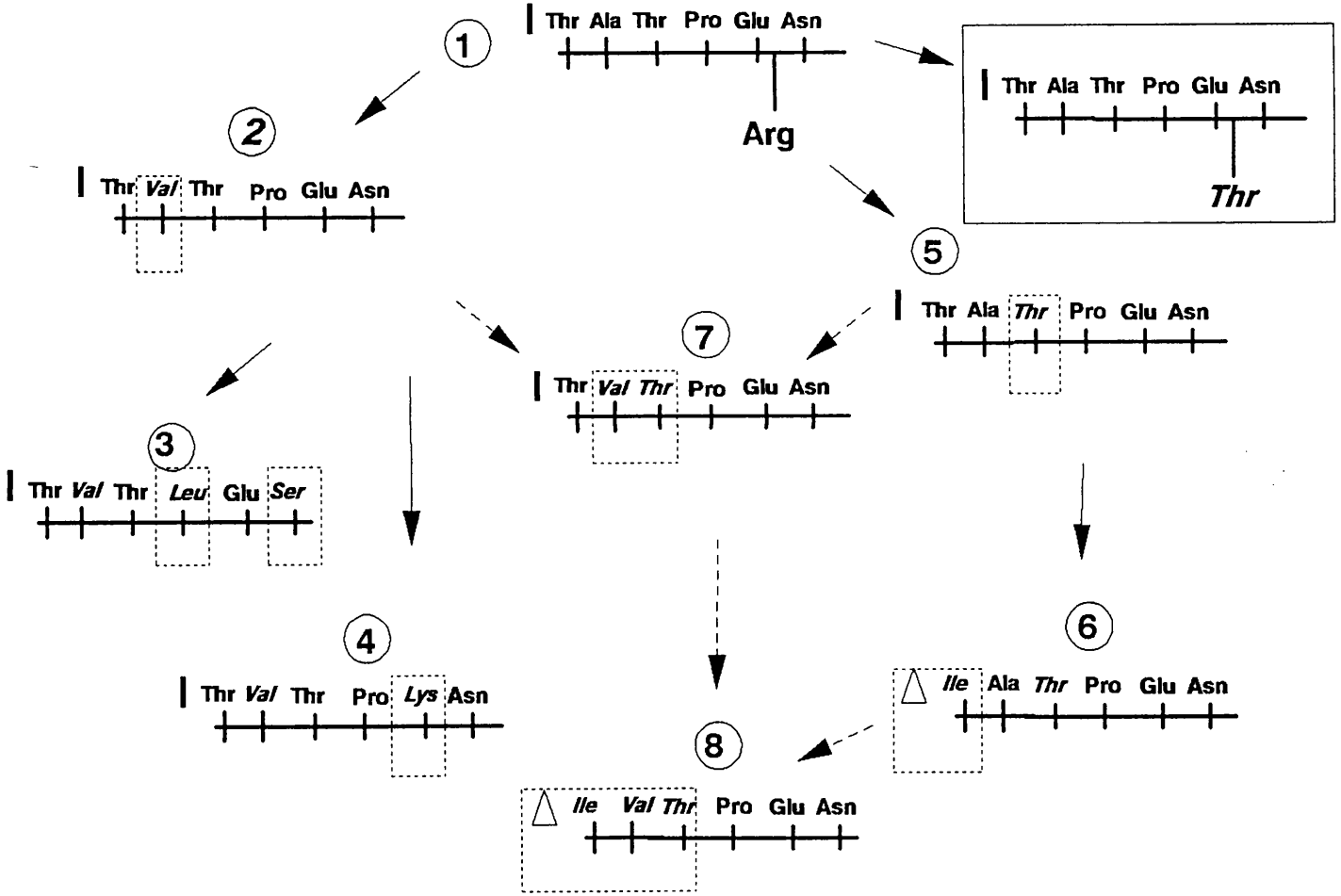


Figure 6.5.2 Phylogenetic tree showing the haplotype carrying the Thr_{4243} variant

For details see Figure 5.3.1

6.6 Discussion

The new variant of the apoB gene generates an Arg to Thr substitution at residue 4243 of mature apoB. Conventional secondary structure protein modelling predicts that this is not a conservative amino acid change - it involves the loss of a positively charged residue. This new polymorphism has an allele frequency of 1-10% in Western populations and no homozygote has yet been identified. The relative rarity of this variant makes statistical analysis of its effect on lipid an lipoprotein levels difficult without a very large population sample. However it can be seen that the effect of this variant is not so great as the effect of Arg_{3500} to Gln which causes FDB-100 (Section 1.4.3). The currently available data does indicate that Thr_{4243} carriers have 4-20% higher serum apoB levels than Arg_{4243} homozygotes, and the magnitude of this effect may thus be comparable to that of the common apoE variants, E2 and E4 on apoB levels (Section 1.2.7). Loss of a charged residue in apoB does not always affect its function: although Arg_{3500} to Gln clearly alters LDL-receptor binding (Innerarity et al. 1988), the Lys_{4154} to Glu substitution does not (Gallagher and Myant 1992), and an association study has failed to find an effect of the Arg_{3611} to Gln substitution on serum lipid levels (Xu et al. 1989). However, it is possible that these charge alterations may have effects other than on receptor-binding that may still be important in atherogenesis. For instance, they may alter the affinity of apoB for arterial proteoglycans (Section 1.2.6).

The existence of this variant provides further evidence for the polymorphic nature of the apoB gene, and the finding that most of the carriers of the Thr_{4243} variant share a common haplotype further emphasises the high levels of linkage disequilibrium within the apoB gene. It remains to be seen how many more variants of apoB there are and what proportion of these have an effect on serum lipid levels or other atherogenic factors.

CHAPTER 7

CLONING AND EXPRESSION OF A PEPTIDE ENCODED BY EXON 29 OF THE APOB GENE.

Summary

A peptide encoded by exon 29 of the apoB gene has been successfully expressed in a baculovirus expression system. The expressed peptide has a size of approximately 60kDa and displays reactivity with the monoclonal antibody $B_{sol}16$ on a Western blot - the antibody recognises an epitope between residues 4157-4189 of mature apoB-100. *In vitro* mutagenesis can be used on the constructs which express this peptide to enable the investigation of the structure/function relationships of the carboxyl-terminal end of apoB.

7.1 Introduction

Evidence has been presented in this thesis that indicates that regions of the apoB gene encoded by exon 29 may have a role in receptor binding. Association studies have also been presented that show that polymorphisms in the carboxyl-terminal region of the protein may generate differences in the affinity of apoB for the receptor. In order to confirm these hypotheses it is necessary to express apoB protein *in vitro* and to investigate the properties of the different variant forms. Expression studies are the only way of confirming that a variant form of a protein truly has a different function and that the detection of an amino acid substitution is not simply a marker for some other functional mutation.

A group of workers in Madison, Wisconsin have developed a method for expressing apolipoprotein genes in a baculovirus expression system using insect tissue culture and the tobacco hornworm *Manduca sexta* larvae. The insect larval fat body is both the target of infection by baculovirus and the site of synthesis of haemolymph lipoproteins making it a useful host for the

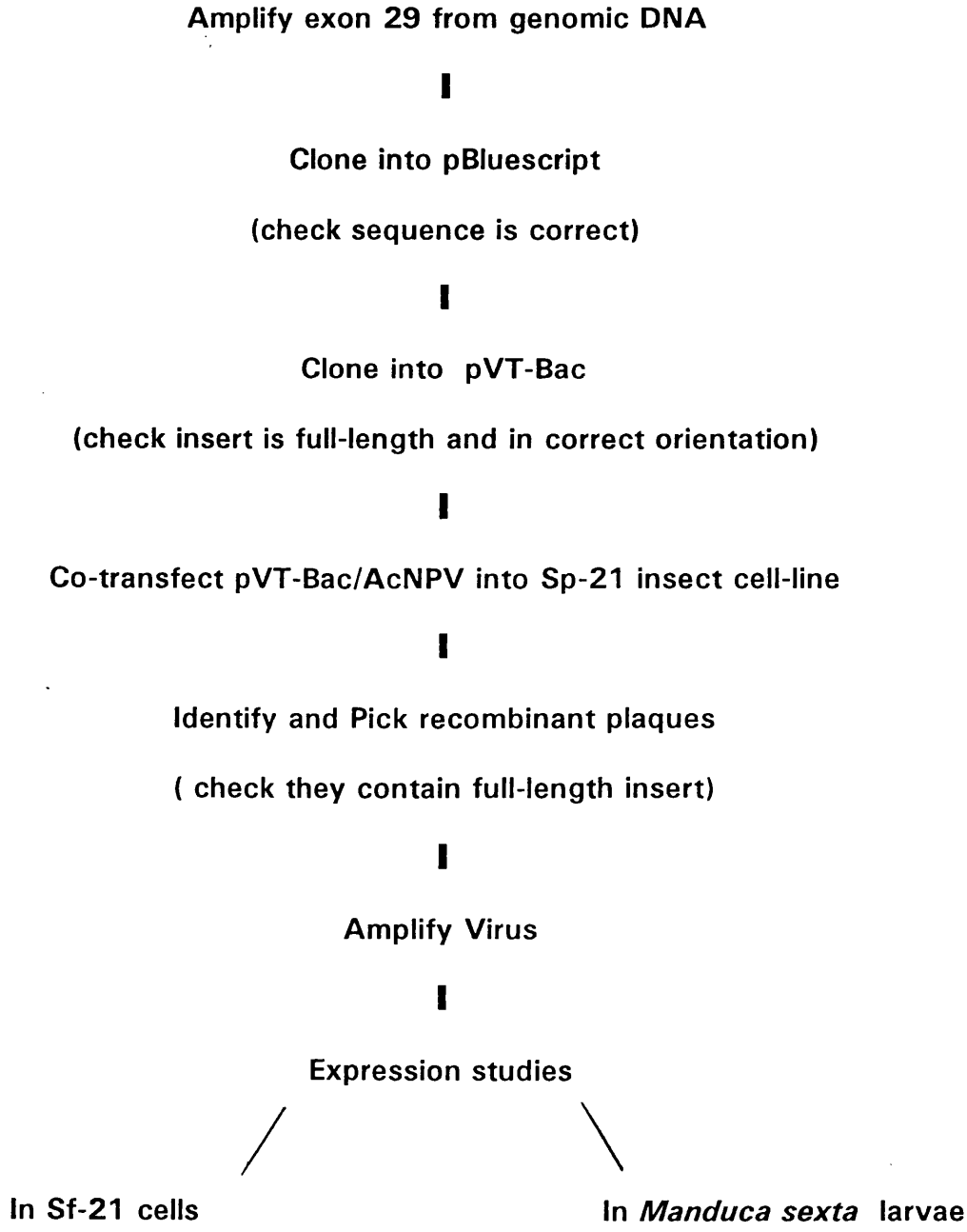
production of recombinant lipoproteins (Gretch et al. 1991). Such lipoproteins can be harvested, at high yield, from the insect haemolymph and separated by preparative ultracentrifugation for use in receptor binding/competition studies. Using this system, these workers have successfully expressed recombinant baculovirus encoding human apoE-3. Approximately one third of the recombinant apoE extracted from the haemolymph of *M. sexta* larvae was lipid associated, in the form of a lipoprotein with a density of < 1.02g/ml. This lipoprotein effectively competed with ^{125}I-LDL for binding to the LDL-receptor on CHO cells. However lipid-free apoE-3 and a second recombinant apoE with an Ala_{152} to Pro mutation, which is a poor ligand for the LDL-receptor (Lalazar et al. 1988), did not compete well. This demonstrates that the *M. sexta* system is capable of expressing recombinant human apoE that retains its native properties.

Since apoB is never found lipid-free *in vivo,* this system should provide a more realistic method of examining the structure and function of apoB than expression in other systems such as *E. coli* or yeast (Sturley et al. 1990) as well as avoiding the complexities of transgenic mice. I have spent three months with the workers in Madison learning the techniques of baculovirus expression. With the help of these colleagues, I constructed and expressed a baculovirus vector carrying apoB exon 29, as a necessary preliminary to future structure/function studies on expressed apoB peptides. A flow diagram of the stages involved in the cloning and expression of the exon 29 construct is shown in Figure 7.1.1

7.2 The cloning and sequencing of an apoB exon 29 insert in pBluescript.

Exon 29 of the human apoB gene was isolated, prior to cloning, by amplification from the genomic DNA of individual 6 from the Sheffield study (Table 2.1.1), generating a 1.9kb fragment of DNA encoding exon 29 of the apoB gene.

Figure 7.1.1 Flow diagram showing the stages in expression of the apoB exon 29 construct.



After amplification with either Vent polymerase or Taq polymerase (Section 2.4) the resulting fragment was force-cloned into pBluescript M13SK- (Strategene, La Jolla CA, USA) using the BamHI and XmaI restriction sites. Clones containing the exon 29 insert were identified by a combination of blue/white selection and PCR of the lysed bacteria with primers **a/b** internal to exon 29 (Figures 4.1.1 and 7.4.1). Of the white colonies identified, the PCR detection system identified 4/50 clones derived from the Taq polymerase amplification, and 3/20 derived from the Vent amplification that contained apoB exon 29 inserts Plasmid DNA was purified from two clones derived from Taq polymerase and three from Vent polymerase amplification and the entire inserts were sequenced (Section 2.8.1) with the M13 universal primer plus seven internal primers to check for correct ligation at the restriction sites and to check the coding sequence of the insert. The construct and position of the sequencing primers are shown in Figure 7.2.1. and the results of the sequencing are tabulated in Table 7.2.2.

The clones derived from Taq polymerase contained many errors including translocations and deletions and so were discarded (results not shown). This was not unexpected in the light of the results presented in Section 3.5. Vent polymerase is a more recently isolated enzyme from *Thermococcus litoralis,* it is thermostable and has a 3' to 5' proof-reading exonuclease which should overcome the problem of errors introduced during PCR.

Of the three clones generated from inserts amplified with Vent polymerase, one had been blunt end cloned in the wrong orientation, a second had a deleted single adenosine at nucleotide 13177 which could have been a mutation in one of the alleles of the genomic DNA amplified, or could have been an artifact introduced by the amplification procedure. The third clone had an exon 29 coding sequence identical to that published (Cladaras et al. 1986, Ludwig et al. 1987) and so this clone, designated DS72, was used in the expression experiments.

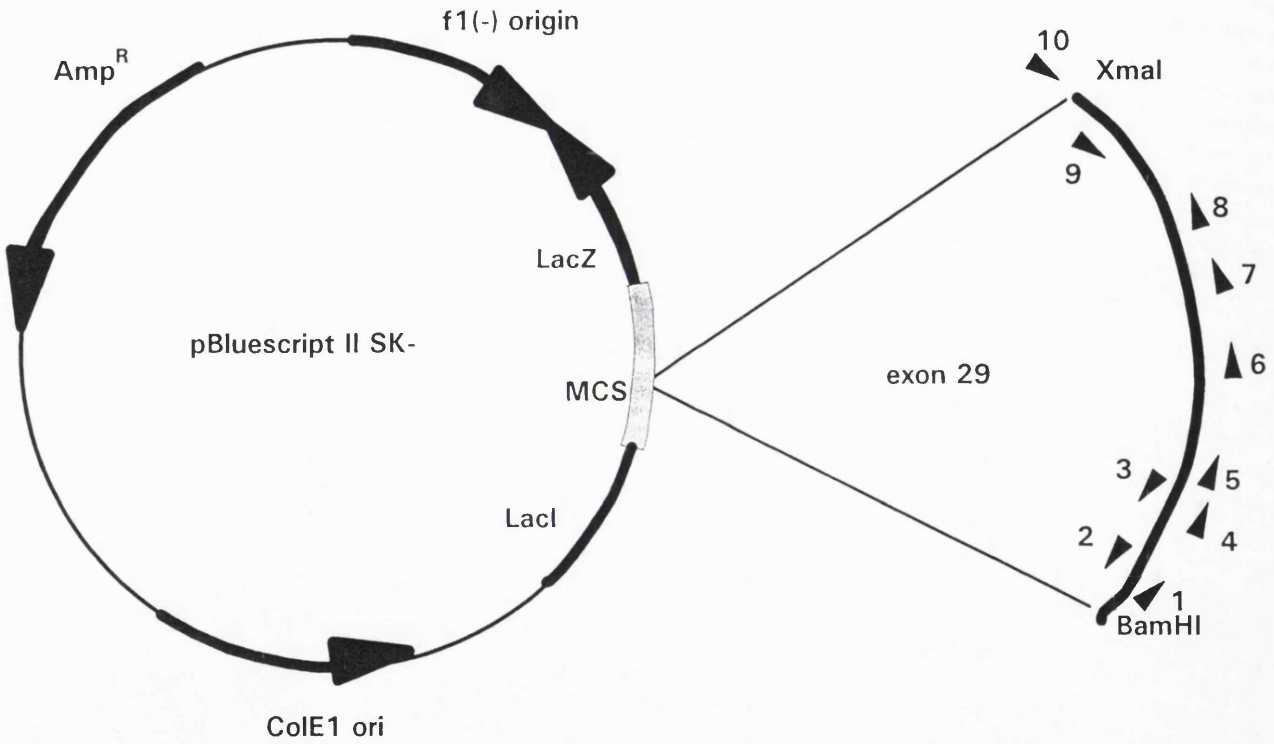


Figure 7.2.1 Position of exon 29 insert cloned in pBluescriptIISK-

The sequences of the primers used to sequence the insert are:- **1** - GCTAACCTCTCTGAAAGACAACTG, **2** - CAGCATTGTTCTGCAGA, **3** - GGAAACTGGAATCTGGGAAGTTC, **4** - CCAGGGAGAGTTCCTAT, **5** - CATAAGGGAGGTAGGGACGGTA, **6** - CCAACTAATAGAAGATAACA, **7** - ATCCAA GTATAGTTGGCT, **8** - CTGCACAGAAATATTCAGG, **9** - TATGGTTTTATCAATATGG, **10** -M13 universal primer. The primers used to amplify the exon 29 insert were:- **5'oligo**: ^{5}cgggatccacAGTCCTCTCCAGATAAAAAACTCACCATATTCand **3' oligo**: ^{5}tcccgGGCTGGCTCACTGTATGGTTT. The bases shown in upper case letters are apoB exon 29 sequence including the (AG) of the 3' intron splice site prior to the first codon (TCC) of exon 29. The bases represented by lower case letters in the **5'oligo** encode a **BamHI** (ggattc) site together with an extra (cg) at the 5' end to facilitate BamHI digestion and an extra (acAG) sequence after the site to put the ligated product in frame with the honey-bee mellitin signal peptide encoded by the pVT-Bac vector. The bases in lower case in the **3'oligo** were added to create a XmaI/SmaI site (cccgGG).

Table 7.2.2 Results of sequencing the Vent amplified clones containing apoB inserts.

Clone	Result
DS71	Blunt-end cloned in inverted orientation, Exon 29 sequence normal.
DS72	Correct orientation. Exon 29 sequence normal.
DS77	Correct orientation. Single deleted adenosine from exon 29 sequence.

7.3 Cloning of the exon 29 insert into pVT-Bac.

The BamHI/XmaI insert from pBluescript M13SK- clone DS72 was next cloned into complementary sites in the baculovirus expression vector pVT-Bac (Figure 7.3.1). Clones containing the insert were identified by PCR using primers **a/b** (Figures 4.1.1 and 7.4.1) The sequence across the cloning sites of the resulting clones was determined to check that the insert was present in the correct orientation and frame with respect to the honey-bee mellitin signal peptide encoded by pVT-Bac. Two clones, termed 72-25 and 72-37, fulfilled the above criteria and plasmid DNA was purified from these on a $CsCl_2$ gradient.

7.4 Co-transfection of constructs into insect cell-lines and detection of recombinant plaques.

Purified DNA from clones 72-25 and 72-37 were used with wild-type *Autographa californica* nuclear polyhedrosis virus (AcNPV) to co-transfect *Sopodoptera frugiperda* Sf-21 insect cells in culture using a cationic liposome mediated transfection procedure (Invitrogen, San Diego, CA USA, Section 2.11.2). Five recombinant viral plaques derived from each construct were picked.

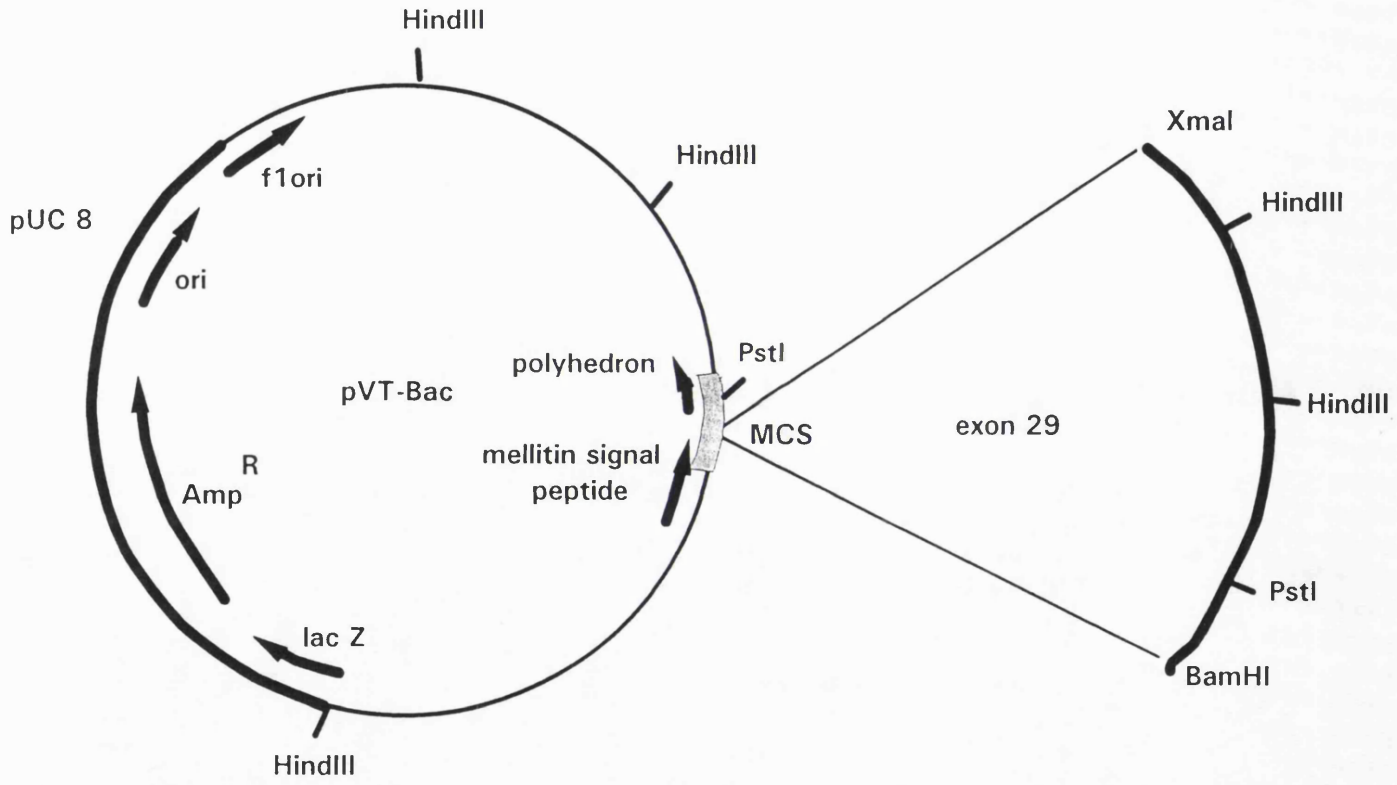


Figure 7.3.1 Position of exon 29 insert in pVT-Bac

The position of the insert in the vector is shown, together with the PstI and HindIII restriction sites.

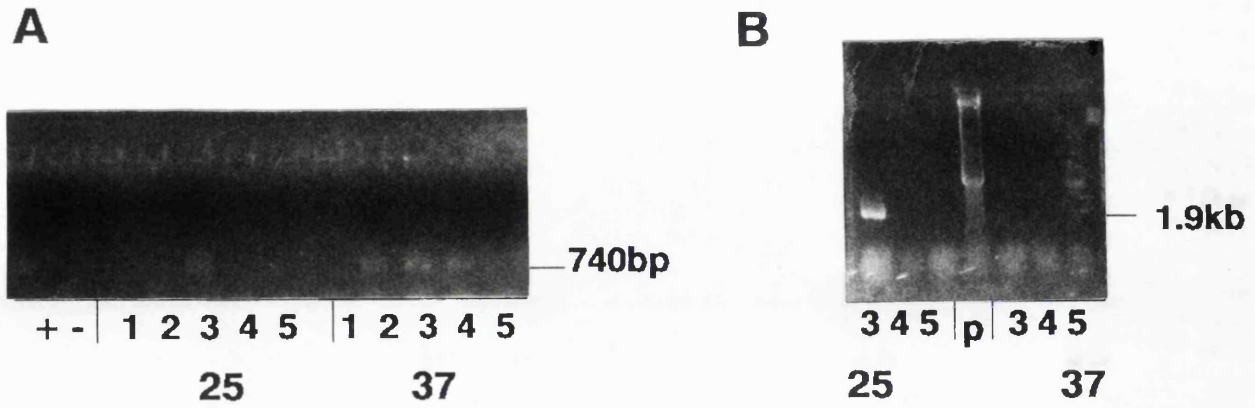


Figure 7.4.1 PCR detection of exon 29 inserts in isolated recombinant virus

A PCR using the primers c/d internal to exon 29, (Figure 4.1.1) which generate a 740bp fragment. The positive control (+) is the pBluescript clone containing the exon 29 insert. The negative control is pBluescript IISK- with no insert. Five plaques derived from each construct: 72-25 (**25**) and 72-37 (**37**) were tested, plaque 72-25-**3** and plaques 72-37-**2,3,4** contain this fragment of exon 29. **B** PCR using the primers Bac1 and Bac2 flanking the cloning sites in pVT-Bac. 5'primer **BAC1**: $^{5'}$CAACGTTGCCCTTGTTTTTATGGTCG, 3'primer **BAC2**: $^{5'}$GAATTTCCTCCTCTTCAGAGAGTCG. Three plaques (**3,4,5**) derived from each construct: **25** and **37** were tested. The pVT-Bac clone containing no insert was used as a negative control (**p**). Only plaque 72-25-3 generates a full length 1.9kb fragment with these primers.

Virus containing apoB exon 29 sequence was again detected by PCR using primers internal to exon 29 and primers flanking the cloning sites in PVT-Bac. The results are shown in Figure 7.4.1. One plaque, derived from clone 72-25 and called 72-25-3, generated fragments of the correct size in both amplification reactions. Three plaques derived from clone 72-37 gave positive results with the primers internal to exon 29 but not with the flanking primers, suggesting that the flanking sequences may have been lost during co-transfection. Work was continued on plaque 72-25-3. After plaque purification and two passes of viral amplification the viral titre was determined to be $4x10^4$ pfu/ml for clone 72-25-3. This viral titre is very low, a titre in the order of 10^5 - 10^6 pfu/ml was expected. The problem has subsequently been found to be associated with a particular passage number of Sf-21 cells, and has been corrected by using cells of a different passage number.

7.5 Detection of expressed peptide from recombinant virus strains.

To determine whether peptide was expressed by the clone, $4x10^6$ Sf-21 cells were infected with 10^4 pfus of virus 72-25-3. Cells infected with 10^6 pfus of apoE-3 expressing baculovirus (Gretch et al. 1991) were used as a positive control for expression and negative control for apoB epitopes on Western blots. Epitopes encoded by exon 29 of the apoB gene were detected on Western blots using the monoclonal antibodies: $B_{sol}7$ which detects an epitope within residues 4521-4536; $B_{sol}16$ (epitope 4157-4189) and MB43 (epitope 4027-4081) (Milne et al. 1989). [The antibodies were kind gifts from Drs. Ross Milne and Linda Curtiss.] Using a cocktail of all three antibodies, a peptide of approximately 60kDa was detected in the media of clone 72-25-3 (Figure 7.5.1A) This was first detectable 72 hours post-infection and increased with time up to 108 hours. Although a 34kDa band, corresponding to the apoE-3 peptide, was visible on the stained blot the antibody cocktail detected nothing in tracks containing this apoE peptide. This demonstrates that the monoclonal antibodies reacted specifically with the apoB epitopes.

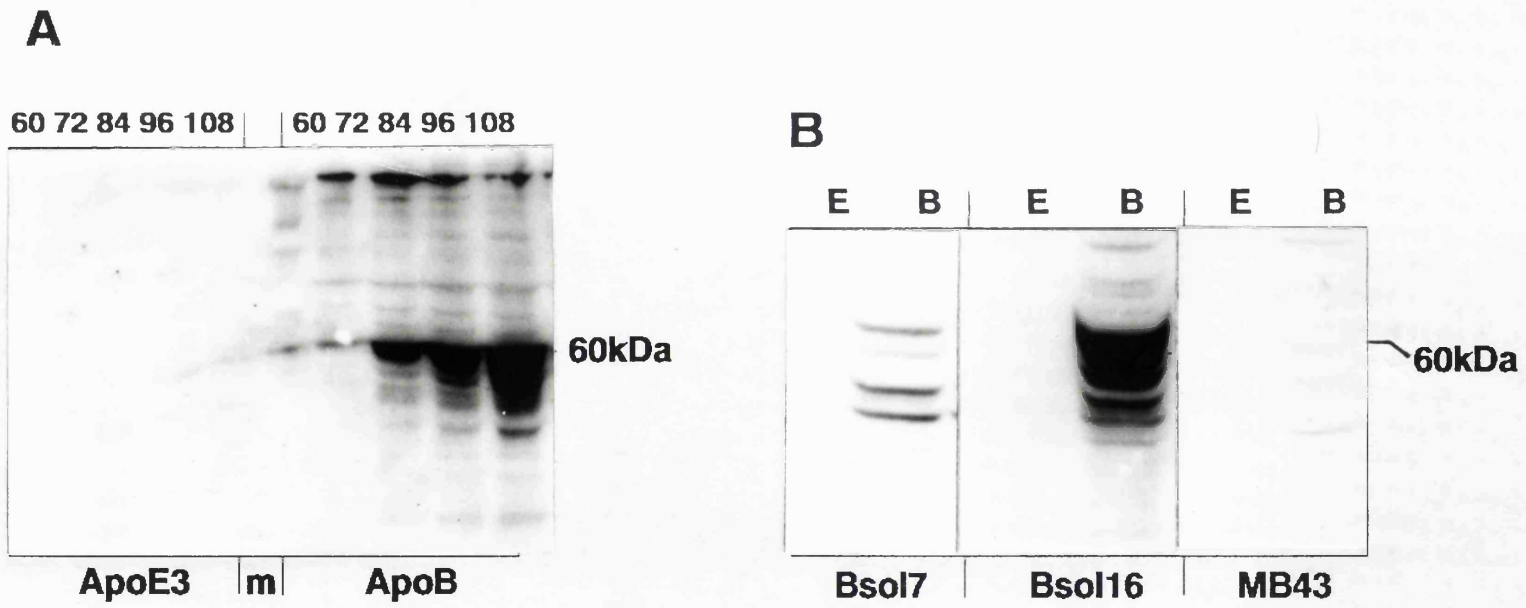


Figure 7.5.1 Western Blot demonstrating peptide expression in Sf-21 cells

A Virus stock from plaque **25-3** was tested using **apoE** virus stock as an expression control. Each track represents a time point in the study: **72,84,96**, and **108** hours. The lane marked **m** is a standard size marker (Bio-rad, Hemel Hempstead, UK) and the sizes of the standards are given on the left. Peptide was detected with a cocktail of $B_{sol}7$, $B_{sol}16$ and MB43. B Peptide from the 108 hour time point from each virus stock: apoE and 72-**25** was tested for reactivity against antibodies: **$B_{sol}7$, $B_{sol}16$ and MB43**. Most of the reactivity seen in **A** is due to $B_{sol}16$.

Media from the 108 hour time-point was then examined for reactivity with each of the monoclonal antibodies alone (Figure 7.5.1b). The reactivity of peptide 72-25-3 with the cocktail can be seen to be mostly due to antibody $B_{sol}16$. Antibodies Bsol 7 and MB43 reacted weakly with media from the clone, each detecting multiple small fragments which could be degradation products of the full length peptide. The apoE control peptide, as expected, did not react with any of the antibodies to apoB.

7.6 Discussion and future experiments.

Exon 29 of the human apoB gene has been cloned and expressed in a baculovirus expression system. The full-length clone in the correct frame and orientation generates a peptide of approximately 60kDa, which is compatible with the size expected from the 1.9kb DNA fragment encoding 534 amino acids. This peptide cross-reacts strongly with monoclonal antibody $B_{sol}16$ but not with $B_{sol}7$ or MB43, possibly because these epitopes are not present on this peptide in the absence of lipid. However, the epitope for MB43 (residues 4027-4081) is only partially available on LDL that has been bound to the LDL-receptor (Milne et al. 1989) and is thus thought to be involved in receptor interaction. Lack of reactivity of the exon 29 peptide with MB43 may be indicative that this pepide, at least in the lipid-free form, would be a poor competitor of LDL in LDL-receptor binding/competition studies.

This peptide, does however, have the potential to be used to further investigate the properties of the carboxyl-terminal region of human apoB. Initially, it is necessary to investigate whether this peptide, in isolation from the rest of the protein, has affinity for lipid - as is predicted from its amino acid composition. Expression of the peptide in *M. sexta* larvae would answer this question. Recombinant virus is injected directly into the heamolymph of fourth instar larvae since, unlike the wild-type virus, it is cannot infect the larvae through ingestion. Six days

post-infection the insect haemolymph can be harvested and fractionated by preparative ultracentrifugation. If the exon 29 peptide is lipid associated it will be detected in the insect lipoprotein fraction .

Lipoprotein associated peptide, isolated from the insect haemolymph, can then be used to investigate whether exon 29 encodes a peptide with any affinity for the human LDL receptor, as has been predicted. The insect apoB associated lipoprotein could be used to compete with ^{125}I-LDL for a normal LDL-receptor on cells in tissue culture and compared to the amount of ^{125}I-LDL bound with no competitor present - a modification of the method of Beisiegel et al. (1981), described in Gretch et al. (1991).

If the exon 29 encoded peptide does have receptor binding activity in isolation from the rest of apoB, including the primary receptor binding domain, *in vitro* mutagenesis techniques could be used to alter regions of the peptide and further investigate their functions. For instance, residues 4003 to 4019, which Chapman et al. (1990) found to have a receptor binding ability, could be deleted to test the importance of these residues in the interaction of the total peptide with the LDL-receptor. In addition, the 10 amino acid insertion found in pigs, and thought to enhance the binding of pig apoB to a human receptor (Ebert et al. 1988), could be inserted between the codons encoding residues 4223 and 4224 to investigate whether this does lead to enhanced receptor binding.

If the receptor binding/competition assay is sufficiently sensitive, it would also be possible to create the naturally-occurring human variants encoded by exon 29 and to investigate their effects on receptor-binding. This would be a method of investigating the mechanisms generating the differences seen in the association studies, and would demonstrate whether the changes themselves are responsible for differences or whether they are simply markers for other

individual or combinations of changes in allelic association. When a monoclonal antibody that distinguishes the Ag(x/y) epitopes becomes available, this could be used in conjunction with the expression of exon 29 to investigate whether the Asn_{4311} to Ser polymorphism alone creates the epitope. However, for studies that can investigate the interaction of residues within the tertiary structure of apoB, it will be necessary to express the entire apoB protein, not just isolated peptides. It should be possible to express a full-length apoB mini-gene construct in the baculovirus system and the workers in Wisconsin are currently attempting this project.

CONCLUSIONS

Chapter 8

The results presented in this thesis suggest that the human apoB gene is polymorphic but not highly prone to recombination. The very high degree of linkage disequilibrium and allelic association seen across the entire gene is a direct result of the low frequency of recombination occurring within the gene. However, it is not clear whether the paucity of recombination events is due to chance or selection. If due to selection, the question then arises as to on what aspect of apoB biology do the selection pressures act? Since CAD does not usually become manifest until after reproductive age, selection is unlikely to act through morbidity from CAD. However it is conceivable that particular variants of apoB could confer resistance to famine or endotoxin and these are both mechanisms through which natural selection could act (Section 5.7). It may be possible that only certain combinations of amino acids, distributed along the length of the apoB polypeptide chain, produce functional apoB protein. In addition to the data presented in Chapters 4 and 5, further evidence supporting this idea comes form the study of Purtell and co-workers (1993) who discovered no single amino acid substitution unique to pig apoB haplotypes.

The results presented in Chapter 3 indicate that although the apoB gene is polymorphic, variation does not frequently occur in the primary receptor binding domain. This region appears to be more highly conserved than the rest of the protein, which is presumably a result of selection. Even the substitution that is known to abolish receptor binding: Arg_{3500} to Gln (Section 1.4.3), is outside the originally proposed primary receptor binding domain (Knott et al. 1985) and demonstrates that changes in this region can have a profound effect on apoB function. The failure to find variation in this region of the protein in patients with LDL with demonstrably reduced affinity for the LDL-receptor is evidence that other regions of the apoB gene may moderate receptor binding.

With the exception of mutations which generate early stop codons and cause hypobetalipoproteinaemia (Section 1.3.2), the effects of variation in other regions of the apoB gene appear to be more subtle and difficult to assay. None- the- less, there is evidence that variation outside of the primary receptor binding domain can generate differences in serum apoB levels and consequent changes in serum LDL- and total cholesterol levels. Evidence is presented here that homozygotes for haplotype 3 [encoding the Ag(x) epitope] have 10-15% lower serum apoB levels, and that carriers of haplotypes 7 and 8 [carrying the XbaI X+ allele in combination with Val_{591}] have 5-15% higher serum apoB levels than carriers of the other haplotypes. In addition the Thr_{4243} variant may also be associated with raised apoB levels (Chapters 4, 5 and 6). These effects are comparable with those generated by the common apoE variants (Section 1.2.7), and may account for most of the differences in apoB levels frequently seen associated with the apoB XbaI RFLP (Section 1.5.3). However the mechanisms by which these variants generate differences in apoB and lipid levels remain to be investigated.

As new, faster screening techniques become available, enabling more apoB alleles to be investigated, more mutations in apoB will no doubt be found. Eventually a picture may emerge as to the nature of the effects of substitutions in different regions of the gene, and this may elucidate the functions of different regions of the protein, and how they interact with one-another.

Developing techniques for the expression of apoB containing lipoproteins *in vitro* is important, as this will provide a method for investigating the mechanism of effect of these variants, and for differentiating the functional variations from the non-functional haplotype markers. Only when the molecular mechanisms causing differences in apoB levels have been determined will it become clear which mutations are clinically important. This will enable rational screening techniques and policies to be developed. The understanding of the mechanisms

behind hyperlipidaemia will also help to determine what advice or treatment should be given to those who are identified as carriers of particular variants.

References

Aalto-Setälä K, Tikkanen MJ, Taskinen MR, Nieminen M, Holmberg P, Kontula K (1988) XbaI and c/g polymorphisms of the apolipoprotein B gene locus are associated with serum cholesterol and LDL-cholesterol levels in Finland. Atherosclerosis 74:47-54

Aalto-Setälä K, Gylling H, Helve E, Kovanen P, Miettinen TA, Turtola H, Kontula K (1989) Genetic polymorphism of the apolipoprotein B gene locus influences serum LDL cholesterol level in Familial Hypercholesterolaemia. Hum Genet 82:305-307

Aalto-Setälä K,Fisher EA, Chen X, Chajek-Shaul T, Zechner R et al. (1992) Mechanism of hypertriglyceridaemia in human apoCIII transgenic mice. J Clin Invest. 90:1889-1900

Aburatani H, Matsumoto A, Itoh H, Yamada N, Murase T, Takaku F, Itakura H (1988) A study of DNA polymorphism in the apolipoprotein B gene in a Japanese population [see comments]. Atherosclerosis 72:71-76

Aggerbeck LP, Wetterau JR, Weisgraber KH, Wu CS, Lindgren FT (1988) Human apolipoprotein E3 in aqueous solution. II. Properties of the amino- and carboxyl-terminal domains. J Biol Chem 263:6249-6258

Allison AC, Blumberg BS (1961) An isoprecipitation reaction distinguishing human serum protein types. Lancet 1:634-637

Armstrong VW, Cremer P, Eberle E, Manke A et al. (1986) The association between Lp(a) concentrations and angiographically assessed coronary atherosclerosis. Dependence on serum LDL levels. Atherosclerosis 62: 249-257

Austin MA & Krauss RM (1986) Genetic control of low density lipoprotein subclasses. Lancet ii: 592-595

Austin MA, King MC, Vranzian KM, Neuman B et al. (1988) Inheritance of low density lipoprotein subclass patterns. Results of complex segregation analysis. Am J Hum Genet 43: 838-

Austin M, Brunzell JD, Fitch WL, Krauss RM. (1990a) Inheritance of low density lipoprotein subclass patterns in familial combined hyperlipidemia. Arteriosclerosis : 10: 520-530.

Austin MA, King MC, Vranizan KM & Krauss RM (1990b) Atherogenic lipoprotein phenotype. A proposed genetic marker for coronary heart disease risk. Circulation 82: 495-506

Austin MA, Wijsman E, Guo SW, Krauss RM, Brunzell JD, Deeb S (1991) Lack of evidence for linkage between low-density lipoprotein subclass phenotypes and the apolipoprotein B locus in familial combined hyperlipidemia. Genet Epidemiol 8:287-297

Austin MA, Horowitz H, Wijsman E, Krauss RM, Brunzell J (1992) Bimodality of plasma apolipoprotein B levels in familial combined hyperlipidemia. Atherosclerosis 92:67-77

Ausubel F M, Brent R, Kingston R E, Moore D D, Seidman J G, Smith J A and Struhl K (eds) (1990) Current protocols in molecular biology. Green publishing and Wiley Interscience, New York.

Aviram M, Lund Katz S, Phillips MC, Chait A (1988) The influence of the triglyceride content of low density lipoprotein on the interaction of apolipoprotein B-100 with cells. J Biol Chem 263:16842-16848

Avogaro P, Bittolo BG, Cazzdato G and Quest GB (1979) Are apolipoproteins better dicriminators than lipids for atherosclerosis? Lancet 1: 901-907.

Babirak SP, Iverius PH, Fujimoto WY, Brunzell JD (1989) Detection and characterization of the heterozygote state for lipoprotein lipase deficiency. Arteriosclerosis 9:326-334

Baum CL, Teng BB, Davidson NO (1990) Apolipoprotein B messenger RNA editing in the rat liver. Modulation by fasting and refeeding a high carbohydrate diet. J Biol Chem 265:19263-19270

Beaty TH, Pregner VL, Virgil DG, Lewis B et al. (1992) A genetic model for control

of hypertriglyceridemia and apolipoprotein B levels in the Johns Hopkins colony of St Thomas Hospital rabbits. Genetics 132:1095-1104

Beisiegal U, Schneider WJ, Goldstein JL, Anderson RGW and Brown MS (1981) Monoclonal antibodies to the Low Density Lipoprotein receptoras probes for the study of receptor-mediated endocytosis and teh genetics of familial hypercholesterolaemia. J Biol Chem 256:11923-11931.

Berg K (1963) A new serum type system in man - the Lp(a) system. Acta pathol Microbiol Scand 59:369-382.

Berg K, Hames C, Dahlen G, Frick MH, Krishan I (1976) Genetic variation in serum low density lipoproteins and lipid levels in man. Proc Natl Acad Sci USA 73:937-940

Berg K, Powell LM, Wallis SC, Pease R, Knott TJ, Scott J (1986b) Genetic linkage between the antigenic group (Ag) variation and the apolipoprotein B gene: assignment of the Ag locus. Proc Natl Acad Sci USA 83:7367-7370

Berg K (1987a) Twin studies of coronary heart disease and its risk factors. Acta Genet Med Gemellol Roma 36:439-453

Berg K (1987b) Genetics of coronary heart disease and its risk factors. In: Bock G, Collins GM (eds) Molecular approaches to human polygenic disease (Ciba Foundation Symposium; 130) 1st ed. Wiley & Sons, Chichester, UK. pp14-28

Berliner JA, Haberland ME (1993) The role of oxidised low-density lipoprotein in atherogenesis. Curr Opin Lipidol 4:373-381

Biemer JJ and Mc Cammon RE (1975) The genetic relationship of abetalipoproteinaemia and hypobetalipoproteinaemia : a report of the occurence of both diseases within the same family. J Lab Clin Med 85: 556-565.

Bersot TP, Russel SJ, Thatcher SR, Pomernacki NK, Mayley RW et al. (1993) A unique haplotype of the apolipoprotein B-100 allele associated with familial defective apolipoprotein B-

100 in a Chinese man discovered during a study of the prevalence of this disorder. J.Lipid Res 34: 1149-1154

Blackhart BD, Ludwig EH, Pieratti VR, Caiati L, Onasch MA, Wallis SC, Powell L, et al (1986) Structure of the human apolipoprotein B gene. J Biol Chem 261:15364-15367

Boerwinkle E, Xiong WJ, Fourest E, Chan L (1989a) Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: application to the apolipoprotein B 3' hypervariable region. Proc Natl Acad Sci U S A 86:212-216

Boerwinkle E, Chan L (1989b) A three codon insertion/deletion polymorphism in the signal peptide region of the human apolipoprotein B (APOB) gene directly typed by the polymerase chain reaction. Nucleic Acids Res 17:4003

Boerwinkle E, Lee SS, Bütler R, Schumaker VN, Chan L (1990) Rapid typing of apolipoprotein B DNA polymorphisms by DNA amplification. Atherosclerosis 81:225-232

Boerwinkle E, Leffert CC, Lin J, Lackner C, Chiesa G, Hobbs HH (1992) Apolipoprotein(a) gene accounts for greater than 90% of the variation in plasma lipoprotein(a) concentrations. J Clin Invest 90:52-60

Boguski MS, Freeman M, Elshourbagy NA, Taylor JM, and Gordon JI. (1986) On computer assiated analysis of biological sequences; proline punctuation, consensus sequences and apolipoprotein repeats. J Lipid Res 27: 1000-34.

Borchardt RA, Davis RA (1987) Intrahepatic assembly of very low density lipoproteins. Rate of transport out of the endoplasmic reticulum determines rate of secretion. J Biol Chem 262:16394-16402

Bostrom K, Boren J, Wettesten M, Sjoberg A, Bondjers G, Wiklund O, Carlsson P, Olofsson SO (1988) Studies on the assembly of apo B-100-containing lipoproteins in HepG2 cells. J Biol Chem 263:4434-4442

Bouthillier D, Sing CF & Davignon J (1983) Apolipoprotein E phenotyping with a single

gel method - application to the study of informative matings. J Lipid Res 24: 1060-1069

Brown WV, Baginsky ML. Inhibition of lipoprotein lipase by an apolipoprotein of human very low density lipoprotein. Biochem Biophys Res Commun 1972: 46: 375-382.

Brown MS and Goldstein JL (1986) A receptor mediated pathway for cholesterol homeostasis. Science 232:34-47.

Bradley WA, Hwang SL, Karlin JB, Lin AH, Prasad SC, Gotto AM and Gianturco SH (1984) Low density lipoprotein receptor binding determinants switch from apolipoprotein E to apolipoprotein B during conversion of hypertriglyceridaemic VLDL to LDL. J. Biol. Chem 259: 14728-35.

Breguet G, Bütler R, Bütler-Brunner E, Sanchez-Mazas A (1990) A worldwide population study of the Ag-system haplotypes, a genetic polymorphism of human low-density lipoprotein. Am J Hum Genet 46:502-517

Brunzell JD, Albers JJ, Chait A, Grundy SM, Groszek E, Mcdonald GB. Plasma lipoproteins in familial combined hyperlipidemia and monogenic familial hypertriglyceridemia. J Lipid Res 1983: 24: 147-155.

Butler R, Brunner E & Morganti G (1974) Contribution to the inheritance of the Ag groups. Vox Sang 26: 485-496

Camejo G, Olofsson S-O, Lopez F, Carlsson P & Benoljers G (1988) Identification of apoB-100 segments mediating the interaction of low density lipoproteins with arterial proteoglycans. Arteriosclerosis 8: 4: 368-377

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution [see comments]. Nature 325:31-36

Carlsson P, Darnfors C, Olofsson SO, Bjursell G (1986) Analysis of the human apolipoprotein B gene; complete structure of the B-74 region. Gene 49:29-51

Chakravarti A, Buetow KH, Antonarakis SE, Waber PG, Boehm CD, Kazazian HH (1984) Nonuniform recombination within the human ß-globin gene cluster. Am J Hum Genet 36:1239-1258

Chapman M J, Forge P, Rouis M, Guo H and Nigon F. (1989) The LDL receptor binding domain of human apolipoprotein B100: Studies with tryptic and synthetic peptides. Atherosclerosis VIII. Amsterdam, Elsevier. 315-318.

Chappell DA, Fry GL, Waknitz MA, Berns JJ (1991) Ligand size as a determinant for catabolism by the low density lipoprotein (LDL) receptor pathway. A lattice model for LDL binding. J Biol Chem 266:19296-19302

Chatterton JE, Phillips ML, Curtiss LK, Milne RW, Marcel YL, Schumaker VN (1991) Mapping apolipoprotein B on the low density lipoprotein surface by immunoelectron microscopy. J Biol Chem. 266:9:5955-5962

Checovich WJ, Fitch WL, Krauss RM, Smith MP, Rapacz J, Smith CL, Attie AD (1988) Defective catabolism and abnormal composition of low-density lipoproteins from mutant pigs with hypercholesterolemia. Biochemistry 27:1934-1941

Cheng C-F, Oosta GM, Bensadown A & Rosenberg RD. (1981) Binding of lipoprotein lipase to endothelial cells in culture. J Biol Chem 256: 12393-12898

Chiesa G, Hobbs H, Koschinsky ML, Lawn RM et al. (1992) Reconstitution of lipoprotein(a) by infusion of human low density lipoprotein into transgenic mice expressing human apolipoprotein(a). J Biol Chem 267: 24369-24374

Cianflone K, Vu H, Walsh M, Baldo A, Sniderman A (1989a) Metabolic response of Acylation Stimulating Protein to an oral fat load. J Lipid Res 30:1727-1733

Cianflone KM, Sniderman AD, Walsh MJ, Vu HT, Gagnon J, Rodriguez MA (1989b) Purification and characterization of acylation stimulating protein. J Biol Chem 264:426-430

Cianflone KM, Maslowska MH, Sniderman AD (1990) Impaired response of fibroblasts

from patients with hyperapobetalipoproteinemia to acylation-stimulating protein. J Clin Invest 85:722-730

Cladaras C, Hadsoupoulou- Cladaras N, et al. (1986) The complete sequence and structural analysis of human apolipoprotein B-100: Relationship between apolipoprotein B-100 and apolipoprotein B-48 forms. EMBO J. 13: 3495-3570.

Cladaras C, Hadzopoulou Cladaras M, Felber BK, Pavlakis G, Zannis VI (1987) The molecular basis of a familial apoE deficiency. An acceptor splice site mutation in the third intron of the deficient apoE gene. J Biol Chem 262:2310-2315

Coleman RD, Kim TW, Gotto AM, Jr., Yang CY (1990) Determination of cysteine on low-density lipoproteins using the fluorescent probe, 5-iodoacetamidofluoresceine. Biochim Biophys Acta 1037:129-132

Colton T (1974) Statistics in Medicine. Little Brown & Co. Boston USA

Converse CA, Keegan WA, Huq L, Series J et al.(1989) Further epidemiological studies on lipid metabolism in retinitis pigmentosa. In: Inherited and environmentally induced retinal degenerations. Liss AR Inc, New York, 39-47

Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak Vance MA (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families [see comments]. Science 261:921-923

Coresh J, Beaty TH, Kwiterovich PO, Jr., Antonarakis SE (1992) Pedigree and sib-pair linkage analysis suggest the apolipoprotein B gene is not the major gene influencing plasma apolipoprotein B levels. Am J Hum Genet 50:1038-1045

Cotton RG, Rodrigues NR, Campbell RD (1988) Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proc Natl Acad Sci USA 85:4397-4401

Cushing GL, Gaubatz JW, Nava ML, Burdick BJ, Bocan TM, Guyton JR, Weilbaecher

D, DeBakey ME, Lawrie GM, Morrisett JD (1989) Quantitation and localization of apolipoproteins [a] and B in coronary artery bypass vein grafts resected at re-operation. Arteriosclerosis 9:593-603

Darnfors C, Nilsson J, Protter AA, Carlsson P, Talmud PJ et al. (1986) RFLPs for the human apolipoprotein B gene: Hinc II & Pvu II. Nucl Acids Res 14: 7135

Davidson NO, Powell LM, Wallis SC, Scott J (1988) Thyroid hormone modulates the introduction of a stop codon in rat liver apolipoprotein B messenger RNA. J Biol Chem 263:13482-13485

Davignon J, Gregg RE, Sing CF (1988) Apolipoprotein E polymorphism and atherosclerosis. Arteriosclerosis 8:1-21

Davis RA, Clinton GM, Borchardt, RA Malone-McNeal et al. (1984) Intrahepatic assembly of very low density lipoproteins. Phosphorylation of small molecular weight apolipoprotein B. J Biol Chem 259: 3383-3386

Davis BH, Kramer RT, Davidson NO (1990) Retinoic acid modulates rat Ito cell proliferation, collagen, and transforming growth factor beta production. J Clin Invest 86:2062-2070

Demmer LA, Levin MS, Elovson J, Reuben MN et al. (1986) Tissue-specific expression and developmental regulations of the rat apolipoprotein B gene. Prot Natl Acad Sci USA 83: 8102-8106

Deeb SS, Failor RA, Brown BG, Brunzell JD, Albers JJ, Motulsky AG, Wijsman E (1992) Association of apolipoprotein B gene variants with plasma apoB and low density lipoprotein cholesterol levels. Hum Genet 88:463-470.

De Loof H, Rosseneu M, Yang C-Y, Li W-H, Gotto AM, Chan L (1987a) Human apolipoprotein B: analysis of internal repeats and homology with other apolipoproteins. J Lipid Res 28:1455-1465

De Loof H, Rosseneu M, Brasseur R, Ruysschaert JM (1987b) Functional differentiation

of amphiphilic helices of the apolipoproteins by hydrophobic moment analysis. Biochim Biophys Acta 911:45-52

Demant T, Houlston RS, Caslake MJ, Series JJ, Shepherd J, Packard CJ, Humphries SE (1988) Catabolic rate of low density lipoprotein is influenced by variation in the apolipoprotein B gene. J Clin Invest 82:797-802

Demant T, Bedford D, Packard CJ, Shepherd J (1991) Influence of apolipoprotein E polymorphism on apolipoprotein B-100 metabolism in normolipemic subjects. J Clin Invest 88:1490-1501

de Knijff P, Kaptein A, Boomsma D, Princen HM, Frants RR, Havekes LM (1991) Apolipoprotein E polymorphism affects plasma levels of lipoprotein(a). Atherosclerosis 90:169-174

Duriez P, Butler R, Tikkanen MJ, Steinmetz J, Vu Dac N, Butler Brunner E, Luyeye I, Bard JM, Puchois P, Fruchart JC (1987a) A monoclonal antibody (BIP 45) detects Ag(c,g) polymorphism of human apolipoprotein B. J Immunol Methods 102:205-212

Duriez P, Vu Dac N, Koffigan M, Puchois P, Demarquilly C, Fievet C, Fievet P, Luyeye I, Bard JM, Fourrier JL, et al (1987b) Detection of human apolipoprotein B polymorphic species with one monoclonal antibody (BIP 45) against low density lipoprotein. Influence of this polymorphism on lipid levels and coronary artery stenosis. Atherosclerosis 66:153-161

Dyer CA, Smith RS, and Curtiss LK (1991) Only multimeres of a syntheic peptide of human apolipoprotein E are biologically active. J Biol Chem 266:23:15009-15015

Dunning AM, Tikkanen MJ, Ehnholm C, Butler R, Humphries SE (1988a) Relationships between DNA and protein polymorphisms of apolipoprotein B. Hum Genet 78:325-329

Dunning, A.M., Talmud, P. and Humphries, S.E. (1988b) Errors in the polymerase chain reaction. Nucl. Acids. Res. 16: 21: 10393.

Dunning AM, Houlston R, Frostegard J, Revill J, Nilsson J, Hamsten A, Talmud P, et

al. (1991) Genetic evidence that the putative receptor binding domain of apolipoprotein B (residues 3130 to 3630) is not the only region of the protein involved in interaction with the low density lipoprotein receptor. Biochim Biophys Acta 1096:231-237

Dunning AM, Renges H-H, Xu C-F, Peacock R, Brasseur R, Laxer G, Tikkanen MJ, Bütler R, Saha N, Hamsten A, Rosseneu M, Talmud P, Humphries S (1992) Two amino acid substitutions in apolipoprotein B are in complete allelic association with the Antigen group (x/y) polymorphism: Evidence for little recombination in the 3' end of the human gene. Am J Hum Genet 50:208-221

Durrington PN, Hunt L, Ishola M, Kane J, Stephens WP (1986) Serum apolipoproteins AI and B and lipoproteins in middle aged men with and without previous myocardial infarction. Br Heart J 56:206-212

Ebert DL, Maeda N, Lowe SW, Hasler Rapacz J, Rapacz J, Attie AD (1988) Primary structure comparison of the proposed low density lipoprotein (LDL) receptor binding domain of human and pig apolipoprotein B: implications for LDL-receptor interactions. J Lipid Res 29:1501-1509

Ehnholm C, Huttunen JK, Rietinen P, Leino M, Mutanen M, Kostiainen E, Pikkarainem J, et al (1982) Effect of a diet on serum lipoproteins in a population with a high risk of coronary disease. N Engl J Med 307:850-855

Ehnholm C, Huttunen JK, Retinen P, Leine LI, Mutanen M, Kostiainen E, Iacono JM, et al (1984) Effect of a diet low in saturated fatty acids on plasma lipid, lipoprotein and HDL subfraction. Arteriosclerosis 4:265-269

Ehnholm C, Mahley RW, Chappell DA, Weisgraber KH et al. (1984b) Role of apolipoprotein E in the lipolytic conversion of beta-very low density lipoproteins to low density lipoproteins in type III hyperlipoproteinemia. Proc Natl Acad Sci USA 81: 5566-5570

Emi M, Hegele RM, Hopkins PN, Wu LL, Plaetke R, Williams RR, Lalouel JM (1991) Effects of three genetic loci in a pedigree with multiple lipoprotein phenotypes. Arterioscler Thromb 11:1349-1355

Fantappie S, Corsini A, Sidoli A, Uboldi P, Granata A, anelli T, Rossi P, Marcovina S, Fumagalli R and Catapano A L (1992) Monoclonal antibodies to human low density lipoprotein identify distinct areas on apolipoprotein B-100 relevant to the low density lipoprotein interaction. J lipid res. 33:1111-1121.

Friedlander Y, Kaufmann NA, Cedar H, Weinberg N, Kark JD (1993a) The role of XbaI polymorphism of the apolipoprotein B gene in determining levels and covariability of lipid and lipoprotein variables in a sample of Israeli offspring with family history of myocardial infarction. Atherosclerosis 98:165-177

Friedlander Y, Kaufmann NA, Cedar H & Kark JD (1993b) Xba I polymorphism of the apolopoprotein B gene and plasma lipid and lipoprotein response to dietary fat and cholesterol: a clinical trial. Clin Genet 43: 223-231

Frostegard J, Hamsten A, Gudland M, Nilsson J (1990) Low density lipoprotein-induced growth of U937 cells - a novel method to determine the receptor binding of low density lipoprotein. J Lipid Res 31,1:33-34

Gabelli C, Baggio G, Bilato S, Martini E et al. (1989) Identification of a new apoB variant associated with hypobetalipoproteinemia. Circulation 80:II-276

Gallagher JJ and Myant NB. (1992) Does the EcoRI polymorphism in the human apolipoprtein B gene affect binding of low density lipoprotein to the low density lipoprotein receptor? Atherosclerosis and Thrombosis:12:2:256-260.

Gandjini H, Gambert P, Athias A, Lallemant C (1991) Resistance to LDL oxidative modifications of an N-terminal apolipoprotein B epitope. Atherosclerosis 89:83-93

Genest J, Sniderman AD, Cianflone K, Teng B et al. (1986) Hyperapobeta-lipoproteinemia. Plasma lipoprotein responses to oral fat load. Arteriosclerosis 6: 297-304

Gibbons GF, Pullinger CR (1987) Regulation of hepatic very-low-density lipoprotein secretion in rats fed on a diet high in unsaturated fat. Biochem J 243:487-492

Ginsberg, HN, Le NA, Goldberg IJ, Gibson JC, Rubinstein A, Wang-Iverson P, Norum R, Brown VW. Apolipoprotein B metabolism in subjects with deficiency of apolipoprotein CIII and AI: Evidence that apolipoprotein CIII inhibits catabolism of triglyceride rich lipoproteins by lipoprotein lipase in vivo. J Clin Invest 1986: 8: 1287-1295.

Gilbert W (1985) Genes-in-pieces revisited. Science 228:823-8

Goate A, Chartier Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L, et al (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease [see comments]. Nature 349:704-706

Goldberg IJ, Scheraldi CA, Yacoub LK, Saxena U, Bisgaier CL (1990) Lipoprotein ApoC-II activation of lipoprotein lipase. Modulation by apolipoprotein A-IV. J Biol Chem 265:4266-4272

Goldstein JL, Schrott HG, Hazzard WR, Bierman EL, Motulsky AG (1973) Hyperlipidemia in coronary heart disease. II. Genetic analysis of lipid levels in 176 families and delineation of a new inherited disorder, combined hyperlipidemia. J Clin Invest 52: 1544-1568.

Goldstein JL, Basu SK, Brunschede GY & Brown MS. (1976) Cell 7: 85-95

Goldstein JL, Brown MS, Anderson RGW < Russell DW, Schnieder WJ (1985) Receptor-mediated endocytosis: conceptrs emerging from the LDL-receptor system. Ann Rev Cell Biol 1:1-39

Graham DL, Knott TJ, Jones TC, Pease RJ, Pullinger CR, Scott J (1991) Carboxyl-terminal truncation of apolipoprotein B results in gradual loss of the ability to form buoyant lipoproteins in cultured human and rat liver cell lines. Biochemistry 30:5616-5621

Gregg RE, Zech LA, Schaefer EJ & Brewer HB Jr. (1981) Type III hyperlipoproteinemia: defective metabolism of an abnormal apolipoprotein E. Science 211: 584-585

Gregg RE, Zech LA, Schaefer EJ Stark D et al. (1986) Abnormal in vivo metabolism of apolipoprotein E4 in humans. J Clin Invest 78: 815-821

Gretch DG, Sturley SL, Friesen PD, Beckage NE andAtte AD (1991) Baculovirus-mediated expression of human apolipoprotein-E in Manduca sexta larvae generates particles that bind to the LDL-receptor. Proc. Nat. Acad. Sci. 88:8530-8533.

Grundy SM (1986) Cholesterol & coronary heart disease. JAMA 256: (20) 2849-2858.

Grundy SM, Chait A & Brunzell JD (1987) Familial Combined Hyperlipidaemia workshop Arteriosclerosis 7: 203-207

Haberland ME, Fless GM, Scanu AM, Fogelman AM (1992) Malondialdehyde modification of lipoprotein(a) produces avid uptake by human monocyte-macrophages. J Biol Chem 267:4143-4151

Hamsten A, Iselius L, Dahlén G, de Faire U (1986a) Genetic and cultural inheritance of serum lipids, low and high density lipoprotein cholesterol and serum apolipoproteins A-I, A-II and B. Atherosclerosis 60:199-208

Hamsten A, Walldius G, Dahlén G, Johansson B, de Faire U (1986b) Serum lipoproteins and apolipoproteins in young male survivors of myocardial infarction. Atherosclerosis 59:223-235

Hamsten A, Eféndic S, Walldius G, Szamosi A, de Faire U, (1987) Glucose tolerance and insulin response to glucose in nondiabetic young male survivors of myocardial infarction. Am Heart J 113:917-927

Hamsten A, (1988) Apolipoproteins, dyslipoproteinaemia and premature coronary heart disease. Acta Med Scand. 223:389-403.

Hansen PS, Gerdes LU, Klausen IC, Gregersen N, Faergeman O (1993) Polymorphisms in the apolipoprotein B-100 gene contributes to normal variation in plasma lipids in 464 Danish men born in 1948. Hum Genet 91:45-50

Hara H, and Howard BV (1990) Characterisation of LDL binding in TR 715-19 cells. A high capacity , stable line for evaluating the binding of LDL to the human LDL-receptor. Atherosclerosis 83 155-165.

Hardman DA, Protter AA, Chen GC, Schilling JW, Sato KY, Lau K, Yamanaka M, Mikita T, Miller J, Crisp T, et al (1987) Structural comparison of human apolipoproteins B-48 and B-100. Biochemistry 26:5478-5486

Hardman DA, Pullinger CR, Hamilton RL, Kane JP et al. (1991) Molecular and metabolic basis for the metabolic disorder normotriglyceridemic abetalipoproteinemia. J Clin Invest 88: 1722-1729

Harpel PC, Gordon BR, Parker TS (1989) Plasmin catalyzes binding of lipoprotein (a) to immobilized fibrinogen and fibrin. Proc Natl Acad Sci U S A 86:3847-3851

Hasstedt SJ, Wu L, Williams RR (1987) Major locus inheritance of apolipoprotein B in Utah pedigrees. Genet Epidemiol 4:67-76

Hayden MR, Kirk H, Clark C, Frohlich J, Rabkin S, McLeod R, Hewitt J (1987) DNA polymorphisms in and around the Apo-A1-CIII genes and genetic hyperlipidemias. Am J Hum Genet 40:421-430

Higgins JA (1988) Evidence that during very low density lipoprotein assembly in rat hepatocytes most of the triacylglycerol and phospholipid are packaged with apolipoprotein B in the Golgi complex. FEBS Lett 232:405-408

Hoeg JM, Meng MS, Demosky Jr SJ, Ronan R Fairwell t, and Brewer Jt HB (1986) Apolipoprotein B : Covalent acylation by fatty acids. Arteriosclerosis 6: 528A.

Hegele RA, Huang L-S, Herbert PN, Blum CB,Buring JE, Hennekens CH, and Breslow J. (1986) Apolipoprotein B - gene DNA polymorphisms associated with myocardial infarction. N Eng J Med 315;24;1509-1515.

Herz J, Hamann U, Rogne S, Myklebost O, Gausepohl H & Stanley KK (1988) Surface

location and high affinity for calcium of a 500-kd liver membrane protein closely related to the LDL-receptor suggest a physiological role as a lipoprotein receptor. EMBO J 7: 4119-4127

Hollander W (1976). Exp Mol Pathol 25: 106-120.

Houlston RS, Turner PR, Revill J, Lewis B, Humphries SE (1988) The fractional catabolic rate of low density lipoprotein in normal individuals is influenced by variation in the apolipoprotein B gene. Atherosclerosis 71:81-85

Huang, L.H., de Graaf, J. and Breslow, J.L. (1988) J. Lipid. Res. 239: 63-67.

Huang LS, Janne PA, de Graaf J, Cooper M, Deckelbaum RJ, Kayden H, Breslow JL, Decklebaum RJD (1990a) Exclusion of linkage between the human apolipoprotein B gene and abetalipoproteinemia [published erratum appears in Am J Hum Genet 47(1):172]. Am J Hum Genet 46:1141-1148

Huang LS, Ripps ME, Breslow JL (1990b) Molecular basis of five apolipoprotein B gene polymorphisms in noncoding regions. J Lipid Res 31:71-77

Huang LS, Gavish D, Breslow JL (1990c) Sequence polymorphism in the human apo B gene at position 8344. Nucleic Acids Res 18:5922.

Huang LS, Kayden H, Sokol RJ, & Breslow JL (1991) ApoB gene nonsense and splicing mutations in a compound heterozygote for familial hypobetalipoproteinemia. J Lipid Res 32: 1341-1348

Hunt SC, Wu LL, Hopkins PN, Stults BM, Kuida H, Ramirez ME, Lalouel JM, Williams RR (1989) Apolipoprotein, low density lipoprotein subfraction, and insulin associations with familial combined hyperlipidemia. Study of Utah patients with familial dyslipidemic hypertension. Arteriosclerosis 9:335-344

Innerarity TL, Weisgraber KH, Arnold KS, Mahley RW, Krauss RM, Vega GL, Grundy SM (1987) Familial defective apolipoprotein B-100: low density lipoproteins with abnormal receptor binding. Proc Natl Acad Sci U S A 84:6919-6923

Innerarity TL, Balestra ME, Arnold KS, Mayley RM, Vega GL, Grundy SM, Young SG (1988) Isolation of defective receptor binding low density lipoproteins from subjects with familial defective apolipoprotein B-100. Arteriosclerosis 8:551

Innerarity JL, Mahley RW, Weisgraber KH et al. (1990a) Familial defective apolipoprotein B-100: a mutation of apolipoprotein B that causes hypercholesterolaemia. J Lipid Res 31:1337-1349.

Innerarity TL (1990b) Familial hypobetalipoproteinemia and familial defective apolipoprotein B100: genetic disorders associated with apolipoprotein B. Current Opinions in Lipidology 1:104-109

International Collaborative Study Group (1986) Metabolic epidemiology of plasma cholesterol. Lancet ii: 991.

Ito Y, Azrolan N, O'Connell A, Walsh A, Breslow JL (1990) Hypertriglyceridemia as a result of human apo CIII gene expression in transgenic mice. Science 249:790-793

Janero DR, Lane MD (1983) Sequential assembly of VLDL apolipoproteins, triacylglycerol and phosphoglycerides by the intact liver cell. J Biol Chem 258: 14496

Jenner K, Sidoli A, Ball M, Rodriguez JR, Pagani F, et al. (1988) Characterisation of genetic markers in the 3' end of the apoB gene and their use in family and population studies. Atherosclerosis 69: 39-49

Jones T, Rajput-Williams J, Knott TJ, Scott J (1989) An MspI RFLP in the apoB promotor. Nucl Acids Res 17:4 72

Kane JP, Havel RJ (1989) Disorders of the biogenesis and secretion of lipoproteins containing the B apolipoproteins. In: The metabolic basis of inherited disease. 6th edition. Eds.Scriver CR, Beaudet WS, Sly WS & Valle D. McGraw-Hill, New York.1139-1164

Kessling AM, Horsthemke B, Humphries SE. A study of DNA polymorphisms around the human apolipoprotein AI gene in hyperlipidaemic and normal individuals. Clin Genet 1985: 28: 296-306.

KleinmanY, Eisenberg D, Oschry Y, Gavish D, Stein O, Stein Y (1985) Defective metabolism of hypertriglyceridaemic Low density lipoprotein in cultured human skin fibroblasts. Normalisation with bezafibrate therapy. J Clin Invest 75: 1796-1803

1796-1803

Knight BL, Perombelon YF, Soutar AK, Wade DP, Seed M (1991) Catabolism of lipoprotein(a) in familial hypercholesterolaemic subjects. Atherosclerosis 87:227-237

Knott TJ, Rall SC, Innerarity TL, Jacobson SF et al. (1985) Human apolipoprotein B: structure of carboxyl-terminal domains, sites of gene expression and chromosomal location. Science 230: 37

Knott JJ, Pease RJ, Powell LM et al. (1986) Complete protein sequence and identification of the structural domains of human apolipoprotein B: Nature 323: 734-738.

Koduri RK, Austin RK, Curtiss LK, Young SJ, Witiztum JL, Smith RS (1991) Definition of a nonlinear epitope on apolipoprotein B for MB47, a monoclonal antibody that completely blocks the binding of LDL to cellular LDL receptors. Arteriosclerosis and Thrombosis 11:5:1404a

Kostner GM (1976) Lp(a) lipoproteins and the genetic polymorphisms of lipoprotein B. In: Low density lipoproteins. Eds Day CE and Levy RS, Plenum New York 229-269

Kowal RC, Herz J, Weisgraber KH, Mayley RW, Brown MS & Goldstein JL (1990) Opposing effects of apolipoproteins E and C on lipoprotein binding to low density lipoprotein receptor-related protein. J Biol Chem 265: 18: 10771-10779

Krempler F, Kostner GM, Roscher A, Haslauer F et al.(1982) Studies on the role of specific cell surface receptors in the removal of lipoprotein(a) in man. J Clin Invest 71: 1431-1441

Krul ES, Tikkanen MJ, Col TG, Davie JM & Schonfeld G (1985) Roles of apolopoprotein B and E in the cellular binding of very low density lipoproteins. J Clin Invest 75: 361-369

Krul ES, Kinoshita M, Talmud P, Humphries SE, Turner S, Goldberg AC, Cook K, Boerwinkle E, Schonfeld G (1989) Two distinct truncated apolipoprotein B species in a kindred

with hypobetalipoproteinemia. Arteriosclerosis 9:856-868

Krul ES, Parhofer KG, Barrett PHR, Wagner RD et al. (1992) ApoB-75, a truncation of apolipoprotein B associated with familial hypobetalipoproteinemia: genetic and kinetic studies. J Lipid Res 33: 1037-1050

Kuusi T, Ehnholm C, Huttunen JK, Kostiainen E, Pietinen P, Leino U, Uusitalo U, et al (1985) Concentration and composition of serum lipoproteins during a low- fat diet at two levels of polyunsaturated fat. J Lipid Res 26:360-367

Kwiterovich PO Jr, Motevalli M & Miller A (1990) Acylation-stimulatory activity in hyperapobetalipoproteinemic fibroblasts: Enhanced cholesterol esterification with another serum basic protein, BPII. Proc Natl Acad Sci USA 87: 8980-8984

Kwiterovich PO (1993) Gentics and molecular biology of familial combined hyperlipidaemia. Curr OP Lipidol 4:133-143

Lackner KJ, Monge JC, Gregg RE, Hoey JM et al. (1986) Analysis of the apolipoprotein B gene and messenger ribonucleic acid in abeta lipoproteinaemia. J Clin Invest: 1397-1404

Lackner C, Boerwinkle E, Leffert CC, Rahmig T, Hobbs HH (1991) Molecular basis of apolipoprotein (a) isoform size heterogeneity as revealed by pulsed-field gel electrophoresis [published erratum appears in J Clin Invest 1991 Aug;88(2):723]. J Clin Invest 87:2153-2161

Law A, Wallis SC, Powell LM, Pease RJ, Brunt H, Priestley LM, Knott TJ, et al (1986) Common DNA polymorphism within coding sequence of apolipoprotein B gene associated with altered lipid levels. Lancet 1:1301-1303

Law A, Scott J (1990) A cross-species comparison of the apolipoprotein B domain that binds to the LDL receptor. J Lipid Res 31:1109-1120

Lawn RM, Wade DP, Hammer RE, Chiesa G et al. (1992) Atherogenesis in transgenic mice expressing human apolopoprotein(a). Nature 360: 670-672

Leitersdorf E, Friedlander Y, Bard JM, Fruchart JC, Eisenberg S, Stein Y (1991) Diverse effect of ethnicity on plasma lipoprotein[a] levels in heterozygote patients with familial hypercholesterolemia. J Lipid Res 32:1513-1519

Leppert M, Breslow JL, Wu L, Hasstedt S, O'Connell P, Lathrop M, Williams RR, White R, Lalouel JM (1988) Inference of a molecular defect of apolipoprotein B in hypobetalipoproteinemia by linkage analysis in a large kindred. J Clin Invest 82:847-851

Lerch PG, Rickli EE, Lergier W & Gillessen D (1980) Localization of individual lysine-binding regions in human plasminogen and investigations on their complex-forming properties. Eur J Biochem 107: 7-13

Li WH, Tanimura M, Luo CC, Datta S, Chan L (1988) The apolipoprotein multigene family: biosynthesis, structure, structure-function relationships, and evolution. J Lipid Res 29:245-271

Liao W & Floren CH (1992) Endotoxin, lipoproteins and heapatic endocytosis. In: Hepatic endocytosis of lipids and proteins. Eds: Windler E & Greten H, Zuckschwerdt Verlag W Germany

Linton MF, Farese RV, Young SG (1993) Familial hypobetalipoproteinaemia. J Lipid Res 34:521-541

Lowe SW, Checovich WJ, Rapacz J, Attie AD (1988) Defective receptor binding of low density lipoprotein from pigs possessing mutant apolipoprotein B alleles. J Biol Chem 263:15467-15473

Ludwig EH, Blackhart BD, Pierotti VR, Caiati L, Fortier C, Knott T, Scott J, Mahley RW, Levy-Wilson B, McCarthy BJ (1989) DNA sequence of the human apolipoprotein B gene. DNA 6:4:363-372.

Ludwig EH, McCarthy BJ (1990) Haplotype analysis of the human apolipoprotein B mutation associated with familial defective apolipoprotein B100. Am J Hum Genet 47:712-720

Lund-Katz S, Ibdah JA, Letizia JY, Thomas MT, Phillips MC (1988) A 13C NMR characterization of lysine residues in apolipoprotein B and their role in binding to the low density lipoprotein receptor. J Biol Chem 263:13831-13838

Lund-Katz S, Innerarity TL, Curtiss LK Krauss RM & Phillips MC (1989) ^{13}C-NMR evidence that substitution of Gln for Arg 3500 in Familial Defective apoB-100 disrupts the conformation of the receptor binding domain. Arteriosclerosis 9: 5: 715a

Ma Y, Schumaker V, Butler R, Sparkes RS (1987) Two DNA restriction fragment length polymorphisms (RFLPs) associated with Ag (t/z) and Ag (c/g) antigenic sites of human apolipoprotein B. Arteriosclerosis 3:301-305

Ma YH, Wang XB, Butler R, Schumaker VN (1989) Bsp 12861 restriction fragment length polymorphism detects Ag(c/g) locus of human apolipoprotein B in all 17 persons studied. Arteriosclerosis 9:242-246

Maeda N, Ebert DL, Doers TM, Newman M, Hasler Rapacz J, Attie AD, Rapacz J, Smithies O (1988) Molecular genetics of the apolipoprotein B gene in pigs in relation to atherosclerosis. Gene 70:213-229

Maniatis T, Fritsch EF, Sambrook J (eds) (1982) Molecular cloning (a laboratory manual), 1st ed. Cold Spring Harbor Laboratory, USA

Marz W, Ruzicka C, Pohl T, Usadel KH, Gross W (1992) Familial defective apolipoprotein B-100; mild hypercholesterolaemia without atheroscleosis in a homozygous patient. Lancet 340:1362

Mahley RE, Innerarity TL, Rall Jr SC & Weisgraber KH (1984) Plasma lipoproteins: apolipoprotein structure and function. J Lipid Res 25: 1277-1294

Mahley RW (1988) Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. Science 240:622-30

Mahley RW & Rall Jr SC (1989) Type III hyperlipoproteinemia (dysbetalipoproteinemia):

the role of apolipoprotein E in normal and abnormal lipoprotein metabolism. In The metabolic basis of inherited disease Eds:Scriver CR, Beaudet Al, Sly WS & Valle D 1195-1213

Malgaretti N, Acquate F, Magnaghi P, Bruno L et al.(1992) Characterization by yeast artificial chromosome cloning of the linked apolipoprotein(a) and plasminogen genes and identification of the apolipoprotein(a) 5' flanking region.Proc Natl Acad Sci USA 89: 11584-11588

McKeigue PM, Shah B, Marmot MG. Central obesity and insulin resistance underlie high diabetes prevalence and cardiovascular risk in South Asians. Lancet (In press)

McLean JW, Tomlinson JE, Kuang WJ, Eaton DL, Chen EY, Fless GM, Scanu AM, Lawn RM (1987) cDNA sequence of human apolipoprotein(a) is homologous to plasminogen. Nature 330:132-137

Miller GJ (1989) Ethnicity, lipoproteins and haemostatic factors. In: Cruickshank JK, Beevers DG (eds) Ethnic factors in health and disease, 1st ed. Wright, London, pp 280-288

Milne-R; Theolis-R Jr; Maurice-R; Pease-RJ; Weech-PK; Rassart-E; Fruchart-JC; Scott-J; Marcel-YL (1989) The use of monoclonal antibodies to localize the low density lipoprotein receptor-binding domain of apolipoprotein B. J-Biol-Chem. 264(33): 19754-60

Monsalve MV, Young R, Jobsis J, Wiseman SA, Dhamu S, Powell JT, Greenhalgh RM, Humphries SE (1988) DNA polymorphisms of the gene for apolipoprotein B in patients with peripheral arterial disease. Atherosclerosis 70:123-129

Montandon AJ, Green PM, Giannelli F, Bentley DR (1989) Direct detection of point mutations by mismatch analysis: application to haemophilia B. Nucleic Acids Res 17:3347-3358

Moreel-JF; Roizes-G; Evans-AE; Arveiler-D; Cambou-JP; Souriau-C; Parra-HJ; Desmarais-E; Fruchart-JC; Ducimetiere-P; et-al (1992) The polymorphism ApoB/4311 in patients with myocardial infarction and controls: the ECTIM Study. Hum-Genet 89(2): 169-75

Myant NB, Gallagher J, Barbir M, Thompson GR, Wile D, Humphries SE (1989) Restriction fragment length polymorphisms in the apo B gene in relation to coronary artery disease. Atherosclerosis 77:193-201

Myant NB (1990) Cholesterol metabolism. LDL and the LDL receptor. Academic Press Ltd, London UK.

Myant NB, Gallagher JJ, Knight BL, McCarthy SN, Frostegard J, Nilsson J, Hamsten A, Talmud P, Humphries SE (1991) Clinical signs of familial hypercholesterolemia in patients with familial defective apolipoprotein B-100 and normal low density lipoprotein receptor function. Arterioscler Thromb 11:691-703

Myklebost O, Rogne S, Hjermann I, Olaisen B, Prydz H (1990) Association analysis of lipid levels and apolipoprotein restriction fragment length polymorphisms. Hum Genet 86:209-214

Namba Y, Tamonage M, Kawasaki H Ohtomo E, Ikeda K (1991) Apolipoprotein E immunoreactivity in cerebral amyloid deposits and neurofibrillary tangles in Alzheimer's disease. Brain Res 541:163-166.

Namba Y, Tsuchiga H, Ikeda K (1992) Apolipoprotein B reactivity in senile plaques, vascular amyloids and neurofibrillary tangles in the brains of patients with Alzheimer's disease. Neuroscience Letters 134:264-266

Naruszewicz M, Selinger E & Davignon J (1992) Lp(a) is apparently more susceptible to oxidative modification induced by incubation with mononuclear cell cultures or Ca++ than LDL; this results in greatly enhanced uptake by macrophages, which seems to involve the oxidized apo(a) moiety. Metabolism 41: 1215-1224

Navajas M, Laurent AM, Moreel JR, Ragab A et al. (1990) Detection by denaturing gradient gel electrophoresis of a new polymorphism in the apolipoprotein B gene. Hum Genet 86: 91-93

Nikkila EA, and Aro A (1973) Family stud y of serum lipids and lipoproteins in coronary

heart disease. Lancet I:954-958.

Nishina PM, Johnson JP, Naggert JK, Krauss RM (1992) Linkage of atherogenic lipoprotein phenotype to the low density lipoprotein receptor locus on the short arm of chromosome 19. Proc Natl Acad Sci U S A 89:708-712

Olofsson SO, Bjursell G, Bostrom K, Carlsson P, Elovson J, Protter AA, Reuben MA, Bondjers G (1987) Apolipoprotein B: structure, biosynthesis and role in the lipoprotein assembly process. Atherosclerosis 68:1-17

Paabo S, Wilson AC (1988) Polymerase chain reaction reveals cloning artefacts [letter]. Nature 334:387-388

Packard CJ, Boag DE, Clegg R, Bedford D & Shepherd J (1985) Effects of 1,2-cyclohexamide modification on the metabolism of very low density lipoprotein apolipoprotein B: potential role of receptors in intermediate density lipoprotein catabolism. J.Lipid Res 26: 1058-1067

Pairitz G, Davignon J, Maillouz H, Sing CF (1988) Sources of interindividual variation in the quantitative levels of apolipoprotein B in pedigree ascertained through a lipid clinic. Am J Hum Genet 43:311-321

Parthasarathy S, Barnett J (1990) Phospholipase A2 activity of low density lipoprotein: evidence for an intrinsic phospholipase A2 activity of apoprotein B-100. Proc Natl Acad Sci U S A 87:9741-9745

Parthasarathy S, Steinberg D, Witztum JL (1992a) The role of oxidized low-density lipoproteins in the pathogenesis of atherosclerosis. Annu Rev Med 43:219-225

Parthasarathy S & Steinberg D (1992b) Cell-induced oxidation of LDL. Carr Op Lipidol 3: 313-317

Paulweber B, Friedl W, Krempler F, Humphries SE, Sandhofer F (1990) Association of DNA polymorphism at the apolipoprotein B gene locus with coronary heart disease and serum

very low density lipoprotein levels. Arteriosclerosis 10:17-24

Peacock R, Dunning A, Hamsten A, Tornval P, Humphries S, Talmud P (1992a) Apolipoprotein B gene polymorphisms, lipoproteins and coronary atherosclerosis: a study of young myocardial infarction survivors and healthy population based individuals. Atherosclerosis 92:151-164

Peacock RE, Hamsten A, Nilsson-Ehle P, Humphries SE (1992b) Associations between lipoprotein lipase gene polymorphisms and plasma correlations of lipids, lipoproteins and lipase activities in young myocardial infarction survivors and age-matched healthy indivdiuals from Sweden. Atheroscleosis 97: 171-185

Peacock RE, Karpe F, Talmud PJ, Hansten A, Humphries SE (1994) Common varaition in genes coding for apolipoproteins E, B and CIII and lipoprotein lipase modulates postprandial responses: a priliminary study. Clin Genet (in press).

Purtell C, Maeda N, Ebert D L, Kaiser M, Sturley S L, Record V, Grunwald K, Nevin N and Attie A D. (1993) Nucleotide sequence of t he carboxy-terminal 10.6kb of the apolipoprotein B gene from spontaneously hypercholesterolaemic pigs. J Lipid Res (submitted)

Perombelon YF, Gallagher JJ, Myant NB, Soutar AK, Knight BL (1992) Lipoprotein(a) in subjects with familial defective apolipoprotein B100. Atherosclerosis 92:203-212

Powell LM, Wallis SC, Pease RJ, Edwards YH, Knott TJ, Scott J (1987) A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. Cell 50:831-840

Pullinger CR, North JD, Teng BB, Rifici VA, Ronhild de Brito AE, Scott J (1989) The apolipoprotein B gene is constitutively expressed in HepG2 cells: regulation of secretion by oleic acid, albumin, and insulin, and measurement of the mRNA half-life. J Lipid Res 30:1065-1077

Rall SC, Welsgraber KH & Mahley RW (1982) Human apolipoprotein E: The complete amino acid sequence. J Biol Chem 257: 4171-4178

Rall SC, Jr., Newhouse YM, Clarke HR, Weisgraber KH, McCarthy BJ, Mahley RW, Bersot TP (1989) Type III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant. J Clin Invest 83:1095-1101

Rall SC, Jr., Mahley RW (1992) The role of apolipoprotein E genetic variants in lipoprotein disorders. J Intern Med 231:653-659

Rapacz J, Chen L, Bütler-Brunner E, Wu M-J, Hasler-Rapacz J, Bütler R, Schumaker V (1991) Identification of the ancestral haplotype for apolipoprotein B suggests an African origin of Homo sapiens sapiens and traces their subsequent migration to Europe and the Pacific. Proc. Natl. Acad. Sci. USA 88:1403-1406.

Rath M, Niendorf A, Reblin T, Dietel M, Krebber HJ, Beisiegel U (1989) Detection and quantification of lipoprotein(a) in the arterial wall of 107 coronary bypass patients [published erratum appears in Arteriosclerosis 1990 Nov-Dec;10(6):1147]. Arteriosclerosis 9:579-592

Rauh G, Schuster H, Muller B, Schewe S, Keller C, Wolfram G, Zollner N (1990) Genetic evidence from 7 families that the apolipoprotein B gene is not involved in familial combined hyperlipidemia. Atherosclerosis 83:81-87

Reaven GM, Chen YD (1988) Role of abnormal free fatty acid metabolism in the development of non-insulin-dependent diabetes mellitus. Am J Med 85:106-112

Renges HH, Peacock R, Dunning AM, Talmud P, Humphries SE (1992) Genetic relationship between the 3'-VNTR and diallelic apolipoprotein B gene polymorphisms: haplotype analysis in individuals of European and south Asian origin. Ann Hum Genet 56:11-33

Robinson MT, Bütler R, Krauss RM (1991) Association of a genetic polymorphism in human apolipoprotein B-100 with intermediate density lipoprotein concentrations. Clin Genet 39:

Ross R (1986) The pathogenesis of atherosclerosis - an update. New Eng J Med 314:8: 488-499

Ross RS, Gregg RE, Law SW, Monge JC et al. (1988) Homozygous hypobetalipoproteinemia: a disease distinct from abetalipoproteinemia at the molecular level. J Clin Invest 81: 590-595

Rosseneu M, De Loof H, De Meutter J, Ruysschaert JM, Vanloo B, Brasseur R (1990) Identification and computer modelling of functional domains in plasma apolipoproteins. In: Brasseur R (ed) Molecular description of biological membrane components by computer aided conformational analysis Vol II, 1st ed. CRC Press Inc, USA, pp173-190

Russell DW, Brown MS, Goldstein JL (1989) Different combinations of cysteine-rich repeats mediate binding of low density lipoprotein receptor to two different proteins. J Biol Chem 264:21682-21688

Saha N, Tay JS, Humphries SE (1992a) Apolipoprotein B-gene DNA polymorphisms (XbaI and EcoRI), serum lipids, and apolipoproteins in healthy Chinese. Genet Epidemiol 9:1-10

Saha N, Tong MC, Tay JSH, Jeyaseelan K, Humphries SE (1992b) DNA polymorphisms of the apolipoprotein B gene in Chinese coronary artery disease patients. Clin Genet 42:4:164-170

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, et al (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487-491

Schlapfer P, Nydegger T, Butler Brunner E, Morgenthaler JJ, Butler R, Blaser K (1987) Two monoclonal antibodies that discriminate between allelic variants of human low density lipoprotein. Hybridoma 6:575-588

Schuh J, Fairclough GH, Haschemeyer RH (1978) Oxygen mediated heterogeneity of apo-low-density lipoprotein. Proc Natl Acad Sci USA 75:3173

Selby JV, Newman B, Quiroga J, Christian JC, Austin MA, Fabsitz RR (1991) Concordance for dyslipidemic hypertension in male twins. JAMA 265:2079-2084

Series J, Cameron I, Caslake M, Gaffney D, Packard CJ, Shepherd J (1989) The Xba1 polymorphism of the apolipoprotein B gene influences the degradation of low density lipoprotein in vitro. Biochim Biophys Acta 1003:183-188

Sharp D, Blinderman L, Combs KA, Kienzle B et al. (1993) Cloning and gene defects in microsomal triglyceride transfer protein associated with abeta lipoproteinaemia. Nature 365: 65-69

Shepherd, J., Bicker, J.S., Lorimer, A.R., and Packard, C.J. (1979) J. Lipid Res. 20: 999-1006.

Shoulders CC, Ball MJ, Mann JI, Baralle FE. (1986)Genetic marker in apolipoprotein AI/CIII gene complex associated with hypercholesterolaemia. Lancet : ii: 1286.

Sing CF & Davignon J (1985) Role of the apolipoprotein E polymorphism in determing normal plasma lipid and lipoprotein variation. Am J Hum Genet 37: 268-285

Sniderman A, Shapiro S, Marpole D, Malcolm I et al. (1980) The association of coronary atherosclerosis and hyperapobetalipoproteinemia (increased protein but normal cholesterol content in human plasma low density lipoprotein). Proc Natl Acad Sci USA 97: 604-608

Sniderman A, Brown G, Stewart BF & Ciantlane K (1992) From familial combined hyperlipidaemia to hyperapo B: unravelling the overproduction of hepatic apolipoprotein B. Curr Op Lipid 3: 137-142

Snyder ML, Polacek D, Scanu AM, Fless GM (1992) Comparative binding and degradation of lipoprotein(a) and low density lipoprotein by human monocyte-derived macrophages. J Biol Chem 267:339-346

Soria LF, Ludwig EH, Clarke HR, Vega GL, Grundy SM, McCarthy BJ (1989) Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. Proc Natl Acad Sci U S A 86:587-591

Stamler J, Wentworth D & Neaton J (1986) Is the relationship between serum cholesterol

and risk of death from corony heart disease continuous and graded? JAMA 256: 2823-2828

Steinberg D, Grundy SM, Mok HYI, Turner JD et al. (1979) Metabolic studies in an unusual case of asymptomatic familial hypobetalipoproteinemia with hypoalphalipoproteinemia and fasting chylomicronemia. J Clin Invest 64: 292-301

Steinberg D, Parthasarathy S, Carew TE, Khoo JC & Witzum JL (1989) Beyond cholesterol modifications of low density lipoprotein that increase its atherogenicity. New Eng J Med: 320: 14:915-924

Steinbrecker UP (1987) Oxidation of human low density lipoprotein results in derivatization of lysine residues of apolipoprotein B by lipid peroxide decomposition products. J Biol Chem 262:8: 3603-3608

Steinmetz A, Jakobs C, Motzny S, Kaffarnik H (1989) Differential distribution of apolipoprotein E isoforms in human plasma lipoproteins. Arteriosclerosis 9:405-411

Stringer CB, Andrews P (1988) Genetic and fossil evidence for the origin of modern humans. Science 239:1263-1268

Sturley SL, Gretch DG, Culbertson MR, Friesen PD, Beckage NE, and attie AD (1990) Heterologous expression of apoproteins. In Molecular Biology of Atheroscleosis, AD Attie, Editor, Elsevier Publishing Co. USA. pp17-26

Talmud PJ, Barni N, Kessling AM, Carlsson P, Darnfors C, Bjursell G, Galton D, et al (1987) Apolipoprotein B gene variants are involved in the determination of serum cholesterol levels : A study in normo-and hyperlipidaemic individuals. Atherosclerosis 67:81-89

Talmud PJ, Lloyd JK, Muller DP, Collins DR, Scott J, Humphries S (1988) Genetic evidence from two families that the apolipoprotein B gene is not involved in abetalipoproteinemia. J Clin Invest 82:1803-1806

Talmud, P.J., King-Underwood, L., Krul, E., Dunning, A., Schonfeld, G. and Humphries S. E. (1989). J. Lipid Res. 30: 1773-1780.

Talmud PJ, Krul ES, Pessah M, Gay G et al. (1994) Donor splice mutation generates a lipid associated apolipoprotein B27.6 in a patient with homozygous hypobetalipoprotienaemia. J Lipid Res (in press).

Teng B, Thompson GR, Sniderman AD, Forte TM et al. (1983) Composition and distribution of low density lipoprotein fractions in hyperapobetalipoproteinemia. Normolipidemia and familial hypercholesterolemia. Proc Natl Acad Sci USA 80: 6662-6666

Teng B, Sniderman AD, Soutar AK & Thompson GR (1986) Metabolic basis of hyperapobetalipoproteinemia. Turnover of apolipiprotein B in low density lipoprotein and its precursors and subfractions compared with normal and Familial Hypercholestrerolemia. J Clin Invest 77:663-672

Teng B, Black DD, Davidson NO (1990) Apolipoprotein B messenger RNA editing is developmentally regulated in pig small intestine: nucleotide comparison of apolipoprotein B editing regions in five species. Biochem Biophys Res Commun 173:74-80

Tessier DC, Thomas DY, Khouri HE, Laliberte F and Vernet T (1991) Enhanced secretion from insect cells of a foreign protein fused to the honeybee melittin signal peptide. Gene 98:177-183.

Tikkanen MJ, Ehnholm C, Kovanen PT, Butler R, Young SG, Curtiss LK, Witztum JL (1987) Detection of two apolipoprotein B species (apoBc and apoBg) by a monoclonal antibody. Atherosclerosis 65:247-256

Tikkanen MJ, Viikari J, Akerblom HK, Pesonen E (1988) Apolipoprotein B polymorphism and altered apolipoprotein B and low density lipoprotein cholesterol concentrations in Finnish children. Br Med J Clin Res Ed 296:169-170

Tikkanen MJ, Xu C-F, Hamalainen T, Talmud P, Sarna S, Huttunen JK, Pietinen P, et al (1989) XbaI polymorphism of the apolipoprotein B gene influences plasma lipid response to diet intervention. Clin Genet 37:327-334

Tikkanen MJ, Xu CF, Hamalainen T, Talmud P, Sarna S, Huttunen JK, Pietinen P,

Humphries S (1990) XbaI polymorphism of the apolipoprotein B gene influences plasma lipid response to diet intervention. Clin Genet 37:327-334

Tribble DL, Holl LG, Wood PD, Krauss RM (1992) Variations in oxidative susceptibility among six low density lipoprotein subfractions of differing density and particle size. Atherosclerosis 93:189-199

Turner, P.R., Konarska, R., Revill, J., Masana, L., La Ville, A., Jackson, P., Cortese, C., Swan, A.V., and Lewis, B. (1984) Lancet (ii): 663.

Tybjaerg Hansen A, Humphries SE (1992) Familial defective apolipoprotein B-100: a single mutation that causes hypercholesterolemia and premature coronary artery disease. Atherosclerosis 96:91-107

Utterman G, Vogelberg KH, Steinmetz A et al. (1979a) Polymorphism of apolipoprotein E.II. Genetics of hyperlipoproteinemia type III. Clin Genet 15: 37-62

Utterman G, Pruin N & Steinmetz A (1979b) Polymorphism of apolipoprotein E. III. Effect of a single polymorphic gene locus on plasma lipid levels in man. Clin Genet 15: 63-72

Utterman G (1985) Genetic polymorphism in apolipoprotein E - Impact on plasma lipoprotein metabolism. In Diabetes, obesity and hyperlipidemias III. Eds Crepaldi G Tiengo A & Baggio G. Excerpta Medica 1-28

Utermann G (1988a) Apolipoprotein polymorphism and multifactorial hyperlipidaemia. J Inherit Metab Dis 11 Suppl 1:74-86

Utermann G, Kraft HG, Menzel HJ, Hopferwieser T, Seitz C (1988b) Genetics of the quantitative Lp(a) lipoprotein trait. I. Relation of LP(a) glycoprotein phenotypes to Lp(a) lipoprotein concentrations in plasma. Hum Genet 78:41-46

Utermann G, Duba C, Menzel HJ (1988c) Genetics of the quantitative Lp(a) lipoprotein trait. II. Inheritance of Lp(a) glycoprotein phenotypes. Hum Genet 78:47-50

Vega GL & Grundy SM (1986) In vivo evidence for reduced binding of low density lipoprotein to receptors as a cause of primary moderate hypercholesterolemia. J Clin Invest 78: 1410-1414

Vilella E, Balanya J, Masana L, Marsal S, La Ville AE, Turner PR (1992) Low density lipoprotein ligand-receptor interactions in normal healthy individuals characterized by their XbaI apolipoprotein B DNA polymorphism. Atherosclerosis 93:145-153

Wade DP (1993) Lipoprotein (a). Curr Op Lipidol 4:244-249

Walsh MT and Atkinson D (1990) Calorimetric and spectroscopic investigation of the unfolding of human apolipoprotein B. J Lipid Res 31:1051-1061

Wang SR, Pessah M, Infante J, Catala D, Salvat C, Infante R (1988) Lipid and lipoprotein metabolism in Hep G2 cells. Biochim Biophys Acta 961:351-363

Wang XB, Schlapfer P, Ma YH, Butler R, Elovson J, Schumaker VN (1988) Apolipoprotein B: the Ag(a1/d) immunogenetic polymorphism coincides with a T-to-C substitution at nucleotide 1981, creating an Alu I restriction site. Arteriosclerosis 8:429-435

Watts GF, Lewis B, Brunt JN, Lewis ES, Coltart DJ, Smith LD, Mann JI, Swan AV (1992) Effects on coronary artery disease of lipid-lowering diet, or diet plus cholestyramine, in the St Thomas' Atherosclerosis Regression Study (STARS) [see comments]. Lancet 339:563-569

Weisgraber KH, Rall SCJ (1987) Human apolipoprotein B-100 heparin-binding sites. J Biol Chem 262:11097-11103

Weisgraber KH (1990) Apolipoprotein E distribution among human plasma lipoproteins : role of the cysteine- arginine interchange at residue112.

Wetterau JR, Aggerbeck LP, Rall SC, Jr., Weisgraber KH (1988) Human apolipoprotein E3 in aqueous solution. I. Evidence for two structural domains. J Biol Chem 263:6240-6248

Whayne TF, Alaupovic P, Curry MD, Lee ET, Anderson PS, Schechter E (1981) Plasma

apolipoprotein B and VLDL-, LDL- and HDL- cholesterol as risk factors in the development of coronary artery disease in male patients examined by angiography. Atherosclerosis 39:411-424

White AL, Graham DL, LeGros J, Pease RJ, Scott J (1992) Oleate-mediated stimulation of apolipoprotein B secretion from rat hepatoma cells. A function of the ability of apolipoprotein B to direct lipoprotein assembly and escape presecretory degradation. J Biol Chem 267:15657-15664

White AI, Rainwater DL, & Lanford RE (1993) Baboon primary hepotocyte cultures were used to study the intracellular processing of apo(a). The immature protein has an unusually prolonged residence time in the endoplasmic reticulum, and the fully glycosylated mature form also accumulates in a pre-secretory compartment. The data suggest that assembly of Lp(a) may occur extracellularly. J Lipid Res 34: 509-517

Wijsman EM, Motulsky AG, Guo SW, Yang M et al.(1992) Evidence against linkage of Familial Combined Hyperlipidemia to the AI/CIII/AIV gene complex (abstract) Circulation 86:I-410

Williams RR, Hunt SC, Hopkins PN, Stults BM, Wu LL, Hasstedt SJ, Barlow GK, Stephenson SH, Lalouel JM, Kuida H (1988) Familial dyslipidemic hypertension. Evidence from 58 Utah families for a syndrome present in approximately 12% of patients with essential hypertension. JAMA 259:3579-3586

Williams KJ, Petrie KA, Brocia RW, Swenson TL (1991) Lipoprotein lipase modulates net secretory output of apolipoprotein B in vitro. A possible pathophysiologic explanation for familial combined hyperlipidemia. J Clin Invest 88:1300-1306

Williams KJ, Fless GM, Petrie KA, Snyder ML, Brocia RW, Swenson TL (1992a) Mechanisms by which lipoprotein lipase alters cellular metabolism of lipoprotein(a), low density lipoprotein, and nascent lipoproteins. Roles for low density lipoprotein receptors and heparan sulfate proteoglycans. J Biol Chem 267:13284-13292

Williams KJ, Petrie KA, Brocia RW & Swenson TL (1992b) Lipoprotein lipase modulated net secretory output of apolipoprotein B in vitro. A possible pathophysiologic explanation for

Familial Combined Hyperlipidemia. J Clin Invest 88: 1300-1306

Wilson C, Wardell MR, Weisgraber KH, Mahley RW, Agard DA (1991) Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E. Science 252:1817-1822

Windler E, Havel RJ. Inhibitory effects of C apolipoproteins from rats and humans on the uptake of triglyceride - rich lipoproteins and their remnants by the perfused rat liver. J Lipid Res 1985: 26: 556-565.

Wojciechowski AP, Farrall M, Cullen P, Wilson TM, Bayliss JD, Farren B, Griffin BA, Caslake MJ, Packard CJ, Shepherd J, et al (1991) Familial combined hyperlipidaemia linked to the apolipoprotein AI-CII-AIV gene cluster on chromosome 11q23-q24. Nature 349:161-164

Wu MJ, Butler E, Butler R, Schumaker VN (1991) Identification of the base substitution responsible for the Ag(x/y) polymorphism of apolipoprotein B-100. Arterioscler Thromb 11:379-384

Xu CF, Nanjee N, Tikkanen MJ, Huttunen JK, Pietinen P, Butler R, Angelico F, Del Ben M, Mazzarella B, Antonio R, et al (1989) Apolipoprotein B amino acid 3611 substitution from arginine to glutamine creates the Ag (h/i) epitope: the polymorphism is not associated with differences in serum cholesterol and apolipoprotein B levels. Hum Genet 82:322-326

Xu CF, Boerwinkle E, Tikkanen MJ, Huttunen JK, Humphries SE, Talmud PJ (1990a) Genetic variation at the apolipoprotein gene loci contribute to response of plasma lipids to dietary change. Genet Epidemiol 7:261-275

Xu CF, Nanjee MN, Savill J, Talmud PJ, Angelico F, Del Ben M, Antonini R, Mazzarella B, Miller N, Humphries SE (1990b) Variation at the apolipoprotein (apo) AI-CIII-AIV gene cluster and apo B gene loci is associated with lipoprotein and apolipoprotein levels in Italian children. Am J Hum Genet 47:429-439

Xu CF, Tikkanen MJ, Huttunen JK, Pietinen P, Butler R, Humphries S, Talmud P (1990c) Apolipoprotein B signal peptide insertion/deletion polymorphism is associated with Ag

epitopes and involved in the determination of serum triglyceride levels. J Lipid Res 31:1255-1261

Yang CY, Chen SH, Gianturco SH, Bradley WA, Sparrow JT, Tanimura M, Li WH, et al (1986) Sequence, structure, receptor-binding domains and internal repeats of human apolipoprotein B-100. Nature 323:738-742

Yang C-Y, Gu Z-W, Weng S-A, Kim TW, Chen S-H, Pownall HJ, Sharp PM, et al (1989) Structure of apolipoprotein B-100 of human low density lipoproteins. Arteriosclerosis 9,1:96-108

Yang CY, Kim TW, Weng SA, Lee BR, Yang ML, Gotto AMJ (1990) Isolation and characterization of sulfhydryl and disulfide peptides of human apolipoprotein B-100. Proc Natl Acad Sci U S A 87:5523-5527

Yao ZM, Blackhart BD, Linton MF, Taylor SM, Young SG, McCarthy BJ (1991) Expression of carboxyl-terminally truncated forms of human apolipoprotein B in rat hepatoma cells. Evidence that the length of apolipoprotein B has a major effect on the buoyant density of the secreted lipoproteins. J Biol Chem 266:3300-3308

Young SG, Bertics SJ, Curtiss LK, Dubois BW, Witztum JL (1987a) Genetic analysis of a kindred with familial hypobetalipoproteinemia. Evidence for two separate gene defects: one associated with an abnormal apolipoprotein B species, apolipoprotein B-37; and a second associated with low plasma concentrations of apolipoprotein B-100. J Clin Invest 79:1842-1851

Young SG, Bertics SJ, Curtiss LK, Witztum JL (1987b) Characterization of an abnormal species of apolipoprotein B, apolipoprotein B-37, associated with familial hypobetalipoproteinemia. J Clin Invest 79:1831-1841

Young SG, Peralta FP, Dubois BW, Curtiss LK, Boyles JK, Witztum JL (1987c) Lipoprotein B37, a naturally occurring lipoprotein containing the amino-terminal portion of apolipoprotein B100, does not bind to the apolipoprotein B,E (low density lipoprotein) receptor. J Biol Chem 262:16604-16611

Young SG, Bertics SJ, Scott TM, Dubois BW, Beltz WF, Curtiss LK, Witztum JL (1987d) Apolipoprotein B allotypes MB19(1) and MB19(2) in subjects with coronary artery disease and hypercholesterolemia. Arteriosclerosis 7:61-65

Young SG, Hubl ST (1989a) An ApaLI restriction site polymorphism is associated with the MB19 polymorphism in apolipoprotein B. J Lipid Res 30:443-449

Young SG, Hubl ST, Chappell DA, Smith RS, Claiborne F, Snyder SM, Terdiman JF (1989b) Familial hypobetalipoproteinemia associated with a mutant species of apolipoprotein B (B-46). N Engl J Med 320:1604-1610

Zuliani G and Hobbs HH (1990) Tetranucleotide repeat polymorphism in the apolipoprotein B gene Nucl Acids Res:18:14:4229.

Zerba K E, Kessling AM, Davignon J, Sing CF (1991) Genetic structure and the search for genotype-phenotype relationships: an example from disequilibrium in the apo B gene region. Genetics 129:525-533

Zilversmit DB (1973). A proposal linking atherogenesis to the interaction of endothelial lipoprotein lipase with triglyceride rich lipoproteins. Circ Res 33: 633-638.

Papers related to this thesis

Papers:-

Dunning A, Tikkanen M, Ehnholm C, Butler R. and Humphries SE. (1987) Relationships between DNA and protein polymorphisms of apolipoprotein B. Hum. Genet. 78:325-329.

Dunning AM, Duriez P, Vu Dac N, Fruchart JC. and Humphries SE. (1988) Association betweeen epitopes detected by monoclonal antibody BIP-45 and the XbaI polymorphism of apolipoprotein B. Clin. Genet. 33: 181-188.

Dunning A, Talmud P. and Humphries SE. (1988) Errors in the Polymerase Chain Reaction. Nuc.Acids Res. 16: 10393.

Talmud P, King-Underwood L, Krul E, **Dunning A,** Schonfeld G. and Humphries SE. (1989) The molecular basis of truncated forms of apolipoprotein B in a kindred with compound heterozygous hypobetalipoproteinaemia. J. Lipid Res. 30: 1773-1779.

Tybjaerg-Hansen A, Gallagher J, Vincent J, Houlston R, Talmud P, **Dunning AM**, Seed M, Hamsten A, Humphries SE. and Myant NB. (1990) Familial defective apollipoprotein B -100: Detection in United Kingdom and Scandinavia, and clinical characteristics of ten cases. Atherosclerosis 80:235-242.

Dunning AM, Houlston R, Frostegard J, Revill J, Nilsson J, Hamsten A, Talmud P. and Humphries SE. (1991) Genetic evidence that the putative receptor binding domain of apolipoprotein B (residues 3130 -3630) is not the only region of the protein involved in interaction with the low density lipoprotein receptor. Biochim. Biophys. Acta 1096:231-237.

Dunning AM, Renges H-H, Xu C-F, Peacock R, Laxer G, Tikkanen MJ, Butler R, Saha N, Hamsten A, Rosseneu M, Talmud P. and Humphies SE. (1992) Two amino acid substitutions in apolipoprotein B are in complete allelic association with the Antigen group (x/y) polymorphism: Evidence for little recombination in the 3' end of the human gene. Am J Hum Genet 50:208-221.

Talmud PJ, Converse C, Krul E, Huq L, McIlwaine GG, Series JJ, Boyd P, Schonfeld G, **Dunning A**. and Humphries SE. (1992) A novel truncated apolipoprotein B (apoB55) in a patient with Hypobetalipoproteinaemia and autosomal dominant retinitis pigmentosa. Clin. Genet. 42: 62-70.

Peacock R, **Dunning A,** Hamsten A, Tornvall P, Talmud P. and Humphries SE. (1992) Apolipoprotein B gene polymorphisms, lipoproteins and coronary atherosclerosis; a study of young myocardial infarction survivors and healthy population based individuals. Atherosclerosis 92:151-164.

Renges H-H, Peacock R, **Dunning AM**, Talmud P. and Humphries SE. (1992) Genetic relationship between the 3'-VNTR and diallelic apolipoprotein B gene polymorphisms: Haplotype analysis in individuals of European and south Asian origin. Ann. Hum. Genet. 56:11-33.

Dunning AM, Renges H-H, Hamsten A, Talmud P. and Humphries S. (1993) A postulated phylogenetic tree for the human apolipoprotein B gene: Unpredicted haoplotypes are associated with elevated apo B levels. Biochim. et Biophys. Acta 1165:271-278.

Reviews:-

Talmud PJ, Houlston RS, **Dunning A**. and Humphries SE. (1989) Apo B gene variants are involved in determining serum cholesterol levels: Towards identifying these variants. In: Human Apolipoprotein Mutants 2. From Gene Structure to Phenotypic Expression. (Eds. Sirtori, C., Franceschini, G., Brewer H., Assmann, G.) Series A: Life Sciences Vol 167, pp75-80.

Duriez P, **Dunning A,** Vu Dac N, Monard-Herkt F, Parra H, Humphries S. and Fruchart JC. (1989) Apolipoprotein B: Immunological methods for the detection of mutants. In: Human Apolipoprotein Mutants 2. From Gene Structure to Phenotypic Expression. (Eds. Sirtori, C, Franceschini, G., Brewer H., Assmann, G.) Series A: Life Sciences Vol 167, pp241-250.

Humphries SE, **Dunning A,** Tybjaerg-Hansen A. and Talmud P. (1989) The use of

molecular biology techniques to analyse a multifactorial trait: Atherosclerosis as an example. Presented at Nordesk Insulin Meeting on "Polygenic Diseases", Olso, June 1989.

Humphries S, **Dunning A,** Xu C-F. and Talmud P. (1991) Strategies to identify common, functionally important sequence polymorphisms in the apo B gene. In: Lipoproteins and the pathogenesis of atherosclerosis. (Eds: Shepherd J, Packard CJ and Brownlie SM) Elsevier Science Publishers BV. pp21-32.

Talmud PJ, Houlston RS, **Dunning A.** and Humphries SE. (1989) Apo B gene variants are involved in determining serum cholesterol levels: Towards identifying these variants. In: Human Apolipoprotein Mutants 2. From Gene Structure to Phenotypic Expression. (Eds. Sirtori, C., Franceschini, G., Brewer H., Assmann, G.) Series A: Life Sciences Vol 167, pp75-80.

Duriez P, **Dunning A,** Vu Dac N, Monard-Herkt F, Parra H, Humphries S. and Fruchart JC. (1989) Apolipoprotein B: Immunological methods for the detection of mutants. In: Human Apolipoprotein Mutants 2. From Gene Structure to Phenotypic Expression. (Eds. Sirtori, C, Franceschini, G., Brewer H., Assmann, G.) Series A: Life Sciences Vol 167, pp241-250.

Humphries SE, **Dunning A,** Tybjaerg-Hansen A. and Talmud P. (1989) The use of molecular biology techniques to analyse a multifactorial trait: Atherosclerosis as an example. Presented at Nordesk Insulin Meeting on "Polygenic Diseases", Olso, June 1989.

Humphries S, **Dunning A,** Xu C-F. and Talmud P. (1991) Strategies to identify common, functionally important sequence polymorphisms in the apo B gene. In: Lipoproteins and the pathogenesis of atherosclerosis. (Eds: Shepherd J, Packard CJ and Brownlie SM) Elsevier Science Publishers BV. pp21-32.

Humphries S, **Dunning A,** Xu C-F, and Talmud P. (1992) Genetic control of plasma lipid, lipoprotein and apolipoprotein levels: From restriction fragment length polymorphisms to specific mutations. In: Cellular and Molecular Biology of

Atherosclerosis. Argenteuil Symposia Proc. (Ed. Gotto AM) Springer-Verlag UK. pp 121-133.

Humphries S, Paul-Hayase H, Gudnason V, Rosseneu M, **Dunning A.** and Talmud P. (1992) Identification of common genetic polymorphisms that determine plasma levels of LDL and HDL. In: Atherosclerosis IX (Eds: O Strein, S Eisenberg, Y Stein) Proceedings of the Ninth Symposium on Atherosclerosis, Chicago, USA. October 1991. pp 255-260.

Appendix

The raw genotype data for the Swedish patient and control groups.

CONTROLS

NUMBER	SIGPEP	71	591	XbaI	2712	4154	4311
4000	IΔ	TI	AA	++	PP	EE	NN
4001	IΔ	TI	AV	-+	PL	EE	NS
4002	IΔ	TI	AV	-+	PL	EE	NS
4003	II	TT	VV	++	PP	EE	NN
4004	IΔ	TI	AA	-+	PP	EE	NN
4005	ΔΔ	II	AA	++	PP	EE	NN
4006	IΔ	TI	AV	-+	PP	EE	NN
4007	II	TT	AV	-+	PP	EK	NN
4008	II	TT	VV	--	PL	EK	NS
4009	II	TT	VV	--	LL	EE	SS
4010	IΔ	TI	AA	++	PP	EE	NN
4011	II	TT	VV	-+	PP	EK	NN
4012	II	TT	VV	--	LL	EE	SS
4013	II	TT	AV	-+	PL	EE	NS
4014	II	TT	VV	-+	PP	EK	NN
4015	IΔ	TI	AV	-+	PL	EE	NS
4016	IΔ	TI	AV	-+	PL	EE	NS
4017	II	TT	AV	--	PP	EK	NN
4018	II	TT	AV	-+	PL	EE	NS
4019	II	TT	AV	--	PL	EE	NS
4020	IΔ	TT	VV	-+	PL	EE	NS
4021	II	TT	AV	-+	PP	EK	NN
4022	ΔΔ	II	AA	++	PP	EE	NN
4023	IΔ	TI	AV	-+	PP	EE	NN
4024	IΔ	TI	AV	-+	PP	EK	NN
4025	II	TT	VV	--	PL	EK	NS
4026	IΔ	TI	AV	-+	PP	EK	NN
4027	II	TT	AV	--	PP	EE	NN
4028	II	TT	AV	--	PP	EK	NN
4029	IΔ	TI	AV	++	PP	EE	NN
4030	IΔ	TI	AV	-+	PL	EE	NS
4031	IΔ	TI	VV	++	PP	EE	NN
4032	ΔΔ	II	AV	++	PP	EE	NN
4033	II	TT	VV	--	PL	EK	NS
4034	II	TT	VV	--	PP	KK	NN
4035	IΔ	TI	AV	++	PP	EE	NN
4036	II	TT	VV	--	PP	KK	NN
4037	II	TT	VV	-+	PP	EK	NN
4038	IΔ	TI	AV	-+	PP	EK	NN
4039	IΔ	TI	AV	-+	PP	EK	NN
4040	ΔΔ	II	AA	++	PP	EE	NN
4041	II	TT	VV	-+	PL	EE	NS
4042	ΔΔ	II	AA	++	PP	EE	NN
4043	ΔΔ	II	AA	++	PP	EE	NN
4044	II	TT	VV	++	PP	EE	NN
4045	IΔ	TI	AA	-+	PP	EE	NN
4046	IΔ	TI	AV	++	PP	EE	NN
4047	IΔ	TI	AV	++	PP	EE	NN
4048	IΔ	TI	AA	++	PP	EE	NN
4049	IΔ	TI	AA	++	PP	EE	NN
4050	II	TT	AV	-+	PP	EK	NN
4051	IΔ	TI	AV	++	PP	EE	NN
4052	IΔ	TI	AV	++	PP	EE	NN
4054	IΔ	TI	AV	--	PL	EE	NS

NUMBER	SIGPEP	71	591	XBAL	2712	4154	4311
4055	IΔ	TI	AA	-+	PP	EE	NN
4056	II	TT	AV	-+	PP	EK	NN
4057	II	TT	VV	--	PP	EK	NN
4058	IΔ	TI	AA	-+	PP	EE	NN
4059	IΔ	TI	AA	++	PP	EE	NN
4060	IΔ	TI	AA	-+	PP	EE	NN
4061	IΔ	TI	AV	++	PP	EE	NN
4062	II	TT	AA	++	PP	EE	NN
4064	IΔ	TI	AV	-+	PL	EE	NS
4065	II	TT	AV	-+	PP	EK	NN
4067	IΔ	TI	AV	-+	PP	EK	NN
4068	IΔ	TI	AA	++	PP	EE	NN
4069	IΔ	TT	VV	--	PL	EK	NS
4070	II	TT	AV	-+	PP	EE	NN
4071	IΔ	TI	AV	-+	PL	EE	NS
4074	IΔ	TI	AV	-+	PP	EK	NN
4076	II	TT	VV	-+	PL	EE	NS
4077	IΔ	TI	AV	-+	PP	EK	NN
4078	II	TT	VV	--	PP	KK	NN
4079	II	TT	AV	-+	PP	EK	NN
4080	II	TT	VV	--	LL	EE	SS
4081	II	TT	AV	++	PP	EE	NN
4082	II	TT	VV	--	LL	EE	SS
4083	II	TT	VV	-+	PP	EK	NN
4084	II	TT	AV	-+	PP	EK	NN
4085	IΔ	TI	AV	-+	PL	EE	NS
4086	II	TT	VV	-+	PL	EE	NS
4087	IΔ	TI	AV	++	PP	EE	NN
4088	II	TT	VV	--	PL	EK	NS
4089	II	TT	VV	--	LL	EE	SS
4090	IΔ	TI	AA	++	PP	EE	NN
4091	ΔΔ	II	AV	++	PP	EE	NN
4092	II	TT	AV	-+	PL	EE	NS
4093	IΔ	TI	VV	-+	PL	EE	NS
4094	II	TT	VV	++	PP	EE	NN
4095	II	TT	AV	-+	PP	EK	NN

PATIENTS

NUMBER	SIGPEP	71	591	XbaI	2712	4154	4311
2002	ΔΔ	II	AA	++	PP	EE	NN
2003	ΔΔ	II	AA	++	PP	EE	NN
2004	ΔΔ	II	AA	++	PP	EE	NN
2005	IΔ	TI	AV	+-	PL	EE	NS
2006	IΔ	TI	AV	+-	PP	EK	NN
2008	II	TT	VV	--	PL	EK	NS
2009	IΔ	TI	AV	+-	PP	EK	NN
2010	IΔ	TI	AV	+-	PL	EE	NS
2011	IΔ	TI	AV	+-	PL	EE	NS
2012	ΔΔ	II	AA	++	PP	EE	NN
2013	II	TT	AV	--	PP	EK	NN
2014	IΔ	TI	VV	+-	PL	EE	NS
2015	II	TT	VV	+-	PL	EE	NS
2016	II	TT	AV	++	PP	EE	NN
2017	ΔΔ	TI	AA	++	PP	EE	NN
2018	II	TT	VV	+-	PP	EK	NN
2019	ΔΔ	II	AA	++	PP	EE	NN
2020	IΔ	TI	AV	+-	PL	EE	NS
2021	II	TT	AV	+-	PL	EE	NS
2022	II	TT	AV	--	PL	EE	NS
2023	II	TT	AV	--	PP	EK	NN
2024	II	TT	AV	--	PP	EK	NN
2025	IΔ	TI	AA	+-	PP	EE	NN
2026	IΔ	TI	AA	--	PP	EE	NN
2027	ΔΔ	TI	AV	--	PL	EE	NS
2028	II	TT	AA	--	PP	EE	NN
2029	IΔ	TI	AV	++	PP	EE	NN
2030	ΔΔ	II	AA	++	PP	EE	NN
2031	II	TT	VV	--	PP	KK	NN
2032	IΔ	TI	AV	+-	PL	EE	NS
2033	ΔΔ	II	AA	+-	PP	EE	NN
2034	IΔ	TI	AV	+-	PP	EK	NN
2035	II	TT	AV	++	PP	EE	NN
2036	IΔ	TI	AV	+-	PL	EE	NS
2037	II	TT	VV	--	LL	EE	SS
2038	IΔ	TI	AV	+-	PL	EE	NS
2039	II	TT	AV	--	PL	EE	NS
2040	ΔΔ	II	AA	++	PP	EE	NN
2041	ΔΔ	II	AA	++	PP	EE	NN
2042	II	TT	AA	--	PP	EE	NN
2043	II	TT	AV	+-	PP	EE	NN
2044	II	TT	VV	+-	PP	EK	NN
2045	IΔ	TI	AV	+-	PP	EE	NN
2046	IΔ	TI	AV	+-	PP	EE	NN
2047	IΔ	TI	AV	+-	PL	EE	NS
2048	IΔ	TI	AA	++	PP	EE	NN
2049	ΔΔ	II	AA	++	PP	EE	NN
2050	IΔ	TI	AA	++	PP	EE	NN
2051	II	TT	VV	++	PP	EE	NN
2052	II	TT	AV	++	PP	EE	NN
2053	II	TT	AV	--	PL	EE	NS
2054	IΔ	TI	AV	+-	PL	EE	NS
2055	IΔ	TI	AV	+-	PL	EE	NS
2056	II	TT	VV	--	PL	EK	NS

2057	IΔ	TI	AV	+-	PP	EK	NN
2058	II	TT	AV	--	PL	EE	NS
2059	II	TT	AA	++	PP	EE	NN
2060	II	TT	AV	--	PL	EK	NS
2061	IΔ	TI	AV	+-	PP	EK	NN
2062	II	TT	AA	+-	PP	EE	NN
2063	ΔΔ	II	AA	++	PP	EE	NN
2064	IΔ	TT	VV	--	LL	EE	SS
2065	II	TT	AV	--	PL	EE	NS
2066	ΔΔ	TI	AV	+-	PL	EE	NS
2067	II	TT	AV	--	PP	EK	NN
2068	IΔ	TI	AV	--	PP	EK	NN
2069	II	TT	AV	+-	PP	EE	NN
2070	II	TT	VV	+-	PP	EK	NN
2071	II	TT	AV	+-	PP	EE	NN
2072	II	TT	AV	++	PP	EE	NN
2074	II	TT	AV	+-	PL	EE	NS
2075	IΔ	TI	VV	+-	PP	EK	NN
2076	II	TT	VV	+-	PP	EK	NN
2077	IΔ	TI	AV	+-	PL	EE	NS
2078	II	TT	AV	+-	PL	EE	NS
2079	IΔ	TI	AA	++	PP	EE	NN
2080	ΔΔ	II	AA	++	PP	EE	NN
2081	ΔΔ	II	AV	+-	PP	EK	NN
2082	II	TT	VV	+-	PL	EE	NS
2083	IΔ	TI	AV	+-	PP	EK	NN
2084	IΔ	TI	VV	--	PL	EK	NS
2085	II	TT	AV	+-	PL	EE	NS
2086	ΔΔ	II	AA	++	PP	EE	NN
2087	ΔΔ	II	AA	++	PP	EE	NN
2088	ΔΔ	TI	AA	--	PL	EE	NS
2089	IΔ	TI	AV	++	PP	EE	NN
2090	IΔ	TI	AV	+-	PL	EE	NS
2091	II	TT	AA	++	PP	EE	NN
2092	IΔ	TI	AV	+-	PL	EE	NS
2093	II	TT	AV	+-	PP	EK	NN
2094	II	TT	VV	+-	PP	EK	NN
2095	IΔ	TI	AV	+-	PL	EE	NS
2096	II	TT	VV	--	LL	EE	SS
2097	II	TT	AA	+-	PP	EE	NN
2098	IΔ	TI	AV	+-	PP	EK	NN
2099	IΔ	TI	AV	++	PP	EE	NN
2100	IΔ	TI	AV	+-	PL	EE	NS
2101	II	TT	AA	+-	PP	EE	NN
2102	II	TT	VV	--	PL	EK	NS
2103	IΔ	TI	AV	+-	PL	EE	NS
2104	ΔΔ	II	AA	++	PP	EE	NN
2105	IΔ	TI	VV	--	PP	EK	NN
2106	II	TT	VV	--	PP	KK	NN
2107	IΔ	TI	AV	+-	PL	EE	NS
2109	IΔ	TI	AV	+-	PP	EK	NN
2110	IΔ	TI	AA	--	PP	EE	NN
2111	II	TT	AV	+-	PL	EE	NS
2112	IΔ	TI	AA	++	PP	EE	NN
2113	IΔ	TI	AV	+-	PP	EK	NN
2114	IΔ	TI	AV	+-	PP	EK	NN
2115	II	TT	AV	+-	PP	EK	NN
2116	ΔΔ	II	AA	--	PP	EK	NN

2117	II	TT	VV	+-	PP	EK	NN
2118	II	TT	VV	+-	PL	EE	NS
2119	IΔ	TI	AA	++	PP	EE	NN
2120	II	TT	VV	+-	PP	EK	NN
2121	II	TT	AV	+-	PP	EE	NN
2122	IΔ	TI	AA	++	PP	EE	NN
2123	IΔ	TI	AV	+-	PL	EE	NS
2124	ΔΔ	II	AA	++	PP	EE	NN
2125	ΔΔ	II	AA	++	PP	EE	NN
2126	II	TT	VV	--	PP	KK	NN
2127	II	TT	AV	++	PP	EE	NN
2128	ΔΔ	II	AA	++	PP	EE	NN
2129	IΔ	TI	AV	+-	PP	EK	NN
2130	II	TT	VV	+-	PP	EK	NN
2131	II	TT	VV	++	PP	EE	NN
2132	IΔ	TI	AV	--	PP	EK	NN
2133	ΔΔ	II	AA	+-	PP	EE	NN
2134	IΔ	TI	AV	+-	PL	EE	NS
2135	II	TT	VV	--	PL	EK	NS
2136	IΔ	TI	AV	+-	PL	EE	NS
2137	II	TT	VV	--	LL	EE	SS
2138	II	TT	AV	--	PP	EK	NN
2139	ΔΔ	II	AA	++	PP	EE	NN
2140	IΔ	TI	AA	++	PP	EE	NN
2141	II	TT	AV	+-	PP	EK	NN
2142	IΔ	TI	AV	++	PP	EE	NN
2143	ΔΔ	II	AA	++	PP	EE	NN
2144	IΔ	TI	AA	++	PP	EE	NN
2145	II	TT	AA	+-	PP	EE	NN
2146	II	TT	AV	+-	PP	EK	NN

Volume 16 Number 21 1988 Nucleic Acids Research

Errors in the Polymerase Chain Reaction

Alison M.Dunning, Philippa Talmud and Steve E.Humphries

Charing Cross Sunley Research Centre, Lurgan Avenue, London, W6 8LW, UK
Submitted August 24, 1988

We have used the Polymerase Chain Reaction (PCR) to amplify a 798bp fragment of the gene for human Apolipoprotein B (Apo B), that contains sequences coding for the putative LDl-receptor binding domain.

5µg genomic DNA from 10 individuals was amplified[1] using 30mer oligonucleotides spanning bases 9599-10397 (inclusive) of the Apo B gene. 30 rounds of amplification were carried out using 5U of Taq Polymerase (Anglian Biotech.) per sample, in a buffer containing : 67mM Tris-HCl (pH 8.8), 6.7mM $MgCl_2$, 16.7mM $(NH_4)_2SO_4$, 10mM β mercaptoethanol, 6.7µM disodium EDTA, 4mg/ml BSA, 10% Dimethylsulphoxide and 330µM (each) dATP, dCTP, dGTP, dTTP, under the regime: 2mins @ 95^0C, 1min @ 55^0C, 5mins @ 70^0C. Amplified DNA was digested with EcoRI and ScaI (Anglian Biotech.) and force-cloned into EcoRI/SmaI cut M13 mp10 (Amersham) using standard techniques. At least 10 clones from each subject were purified. Clones were sequenced using the Sequenase kit (USB Inc.) and analysed on 8% denaturing polyacryamide gels.

Initially the sequence of one clone per individual was determined.Out of the total of 8000 bases sequenced (10 individuals), 22 differences were detected (Table). No clone was identical to the published sequences[2]. Since any genuine base change should be present in approximately half the clones analysed (assuming the individual to be heterozygous), we subsequently analysed all 10 clones from each individual. None of the initial differences found were present in any of the other clones, although all of them were reproduced upon resequencing of the original clones. This implies that all the base differences seen were artefacts generated by the PCR.

The most common changes found were from A to G and from T to C. 17/22 (77%) of the changes noted were associated with a run of bases of the same sequence (Table). This may be an indication of the mechanism by which the errors are inserted.

These observations indicate that the interpretation of sequence changes rom cloned, amplified DNA must be made with caution. Direct sequencing of the PCR material would probably overcome these artefacts[3].

Base Change	Sequence
A > C	ATTTAAGT
A > G	TCACAAAT
	CATCAAAT
	TACAAAGC
	GATTGAAG
	GCCACAGC
C > T	CTCCCCAG
	AGGCACAG
G > A	GAGCTGCC
	TGCCAGTC
T > A	AGTTGTCA
	CAGCATGC
	TGCACTGC
T > C	GTTTATCA
	GGACCTTT
	ATGATTTC
	ATTGTTGC
	CTTCATTG
	TGAAGTTA
Lost A	GAAAAAGG
Lost G	TTCCAGTT
Lost T	AAGTTTGA

Acknowledgements: This work was funded by the British Heart Foundation & the Charing Cross Sunley Research Trust.

References

1. Saiki RK, Gelfand DH, et al. (1987) Science 239: 487-491.
2. Olofsson S-O, Bjursell G, et al.(1987) Atherosclerosis 68: 1-17.
3. Paabo S. and Wilson A.C. (1988) Nature 334:387.

Biochimica et Biophysica Acta, 1096 (1991) 231–237

ADONIS 092544399100070Q

BBADIS 61024

Genetic evidence that the putative receptor binding domain of apolipoprotein B (residues 3130 to 3630) is not the only region of the protein involved in interaction with the low density lipoprotein receptor

Alison M. Dunning [1], Richard Houlston [1,*], Johan Frostegård [2], John Revill [3], Jan Nilsson [2], Anders Hamsten [2], Philippa Talmud [1] and Steve Humphries [1]

[1] *Arterial Diseases Research Group, Charing Cross Sunley Research Centre, London (U.K.),* [2] *King Gustaf V Research Institute, Karolinska Institute and Department of Internal Medicine, Karolinska Hospital, Stockholm (Sweden) and* [3] *Lowedges, Sheffield (U.K.)*

(Received 14 August 1990)

Key words: Apolipoprotein B; Receptor binding domain; LDL receptor; Hypercholesterolemia; Gene mutation; (Human fibroblast)

We have searched for sequence differences in the region of the apolipoprotein B (*apo B*) gene encoding amino acids 3130–3630 in eight individuals with reduced affinity of low density lipoprotein (LDL) for the normal LDL-receptor. All individuals were hypercholesterolaemic and were selected either on the basis of reduced fractional catabolic rate (FCR) of autologous LDL or substantially reduced binding of their LDL to normal LDL-receptors determined by an in vitro cell growth assay using the U937 macrophage-like cell line. Segments of the *apo B* gene were amplified by the polymerase chain reaction. Using a combination of cloning and sequencing the amplified fragment, together with chemical cleavage mismatch analysis, no sequence differences were identified in this region of the gene. We therefore conclude that variation outside the region of the *apo B* gene that codes for amino acids 3130–3630 must be responsible for the reduced LDL clearance in these patients.

Introduction

Apolipoprotein B100 (apo B) is a 55 kDa protein that is produced by the liver and secreted as a constituent of very low density lipoprotein (VLDL). During the metabolism of VLDL other apolipoproteins are removed, leaving apo B as the sole protein component of LDL. Its structure serves both to maintain the integrity of the LDL particle [1] and as ligand for the LDL-receptor. The amino acids comprising the region of apo B that interact with the LDL-receptor have been defined by several approaches. Firstly, by homology to the apolipoprotein E (apo E) receptor-binding domain [2], secondly by the ability of certain monoclonal antibodies raised to apo B to block binding of LDL to the LDL-receptor [2] and thirdly by the capability of synthetic peptides to bind to a normal LDL-receptor [1]. Taken together these methods suggest that two regions are involved in receptor binding, one spanning amino acids 3147 to 3157 and a second comprising residues 3345 to 3381 (Fig. 1).

This putative receptor binding region has been further expanded in the light of the discovery and characterisation of familial defective apo B100 (FDB) [3,4]. FDB is a rare, co-dominantly inherited disorder, generated by a substitution of glutamine for arginine at residue 3500 of apo B. This mutation severely reduces binding of LDL-apo B to the LDL-receptor and thus causes moderate to severe hypercholesterolaemia in carriers [5,6,7]. Residue 3500 is 120 amino acids outside the initially-defined receptor-binding domain, but clearly affects the interaction of apo B with the receptor.

It has been shown in population studies, that varia-

* Present address: Department of Paediatric Genetics, Royal Free Hospital Medical School, London (U.K.)

Abbreviations: apo B, apolipoprotein B; LDL, low density lipoprotein; FCR, fractional catabolic rate; VLDL, very low density lipoprotein; apo E, apolipoprotein E; FDB, familial defective apo B100; PCR, polymerase chain reaction.

Correspondence: A.M. Dunning, Arterial Diseases Research Group, Charing Cross Sunley Research Centre, Lurgan Avenue, Hammersmith, London W6 8LW, U.K.

tion at a polymorphic *Xba*I site within the *apo B* gene is associated with differences in serum cholesterol levels, with the X + allele (presence of the *Xba*I cutting site) being associated with raised levels [8–10]. These studies demonstrate that variation at the *apo B* gene locus is involved in the determination of serum cholesterol levels, but the mechanism of this association is unknown. The base change that creates the variant *Xba*I site does not alter an amino acid, as it is in the third (wobble) position of a codon for threonine 2488 [11], and so cannot be the direct cause of the effects seen. Thus we postulate a second sequence difference, elsewhere in the gene and in linkage disequilibrium with the X + allele, which does lead to raised serum cholesterol levels. This sequence difference may, as with the apo B_{3500} mutation, change an amino acid in the protein and alter the affinity of LDL for the receptor. Evidence in favour of this has come from two in vivo studies that have shown that the X + allele is associated with reduced fractional catabolic rates (FCR) of autologous LDL [12,13].

Further confirmation of these results have come from a study that used normal human dermal fibroblasts to compare simultaneously the binding, internalisation and degradation rates of LDL from individuals homozygous for the presence or absence of the *Xba*I cutting site [14]. This study found that, in a group of older, normolipidaemic people as well as in moderately hypercholesterolaemic patients, LDL from individuals homozygous for the X + allele was degraded slower than that from individuals homozygous for the X − allele. However, this difference was small and could only be detected when the two types of LDL were compared directly in the same cell culture dish.

Recently a cell growth assay using the U937 monocyte cell line has been developed that recognises LDL particles with altered affinity for the LDL-receptor [15]. U937 cells have no scavenger receptor and are thus unable to take up modified forms of LDL. In addition, during transformation these cells have lost their ability to synthesise cholesterol by the endogenous synthetic pathway. Thus, under serum-free conditions, growth is dependant upon and proportional to exogenous cholesterol taken up by the LDL-receptor. This assay can thus be used to identify individuals whose LDL binds with reduced affinity to the LDL-receptor.

In this study we have used several different techniques to compare the sequences of the *apo B* gene coding for amino acids 3130–3630 from eight hyperlipidaemic patients. Three of these patients have LDL with substantially reduced affinity for the normal LDL-receptor and the others have a low fractional catabolic rate of autologous LDL. Seven of these eight patients are homozygous for the *Xba*I X + allele, and we anticipate that several of these patients will share the postulated common lipid-raising sequence difference that is in linkage disequilibrium with this allele.

Subjects and Methods

Subjects

The seven patients who are homozygous for the *Xba*I X + allele have been described in previous studies: patients 1–5 from Glasgow [12]; patients 6 and 7 from Sheffield [13]. Patient 8 from Stockholm is a 51-year-old man who suffered myocardial infarctions at the age of 43 and 47 and an iliaco-femoral artery thrombosis at the age of 49. Patients with FH were excluded on the basis of standard criteria [12], from the original studies from which patients 1 to 7 were drawn and, in the case of patient 8, on the grounds of normal LDL-receptor number and function on the patient's cultured lymphocytes. The 'probe' DNA used in the chemical mismatch cleavage assay was derived from a subject from the Glasgow Study with genotype X-X- and high FCR. (individual 9, Table I) [12]. A positive control for the mismatch assay was DNA from a patient heterozygous for the apo B_{3500} substitution (individual 10, Table I).

U937 cell growth assays

Fasting blood samples were taken into tubes containing 1 mg/ml Na_2EDTA and spun at $400 \times g$ at 4°C. The plasma was removed and stored on wet ice whilst being transported to Stockholm. Assays, in triplicate, were carried out within 3 days of blood collection. Assays were performed as described previously [15]. Briefly, cells are incubated in serum-free medium for 24 h, before resuspension in medium containing test LDL and ^{3}H-thymidine for a further 24 h. Cells are then harvested and the amount of incorporated ^{3}H determined in a liquid scintillation spectrometer.

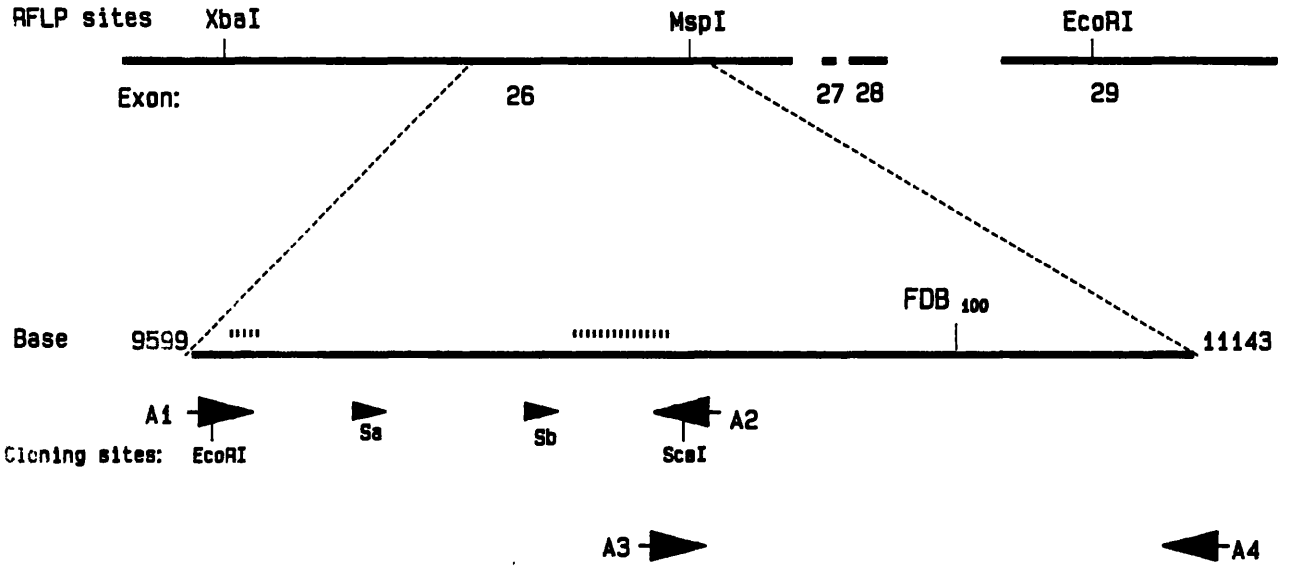
DNA isolation and Southern blot analysis

These were performed using standard techniques as described previously [10].

Polymerase chain reaction (PCR)

Regions of the *apo B* gene were amplified by PCR [16]. The sequences of the oligonucleotide amplimers used and their arrangements are shown in Fig. 1. Oligonucleotides were synthesised on a Pharmacia Gene Assembler (Pharmacia, Uppsala, Sweden) and the protective DMTr group was removed according to manufacturer's instructions. The PCR reactions were performed in an automated thermal cycler, (Cambio, Cambridge, U.K.) using 50 ng genomic DNA, 250 ng each amplimer and *Thermus aquaticus* DNA polymerase (Perkin Elmer/Cetus, CT, U.S.A.) in the manufacturers recommended buffer. After initial denaturing of the DNA for 5 min at 95°C, and annealing of the primers (55°C, 3 min), the programme used was 50 cycles of: 72°C, 3 min; 95°C, 1 min; 55°C, 1 min.

DNA generated by the PCR reactions was purified by Geneclean (Bio101, La Jolla, U.S.A.) according to



Oligo sequences, 5' to 3':-

Amplimers	**A1**	**TTGAAGGAATTCTTGAAAACGACAAAGCAA**	**Sequimers**	**Sa**	**CAAATTCCTGGATACACTGT**
	A2	**CACTTCCATATTTTTCGTGGTTAAGCTCAC**		**Sb**	**GAATACCAATGCTGAACTTT**
	A3	**CATAACAGTACTGTGAGCTTAACCACGAAA**			
	A4	**AAGGATCCTGCAATGTCAAGGTGTGCCTTT**			

Fig. 1. Map of the 3′ end of the *apo B* gene, showing the positions and sequences of oligonucleotides used for PCR and sequencing. The bold dotted lines indicate the nucleotides encoding protein regions with homology to the apo E receptor binding domain (nucleotides 9648–9681 and 10245–10353). The relative positions of base changes creating RFLP sites and the apo B_{3500} mutation are shown.

manufacturers instructions and eluted into TE buffer to give a DNA concentration of 50 to 100 ng/μl.

Cloning and sequencing

The products of the entire PCR reaction were digested with restriction enzymes *Sca*I and *Eco*RI (Boehringer-Mannheim, U.K.) (see Fig. 1) and ligated into complementary sites in phage M13 mp11 (Pharmacia, Uppsala, Sweden) using T4 DNA ligase (BRL, Paisley, Scotland) according to manufacturers instructions. Ligated phage were used to transform competent *E. Coli* strain JM 101 cells and plated out onto medium containing X.gal and IPTG. (Anglian Biotech., U.K.) Single white plaques were picked and the phage DNA isolated and purified ready for sequencing using T7 DNA polymerase (Sequenase, USB, Cleveland, U.S.A.) following the manufacturer's recommended protocol. Sequencing reactions were resolved on 8% denaturing, polyacrylamide gels on a BRL model S2 apparatus (BRL, Paisley, U.K.). Internal primers were used to resolve sequences furthest from the cloning sites, (Fig. 1). It was necessary to analyse 10 clones recovered from amplification of each patient's DNA in order to differentiate between PCR misincorporation errors and heterozygous base changes [17].

Chemical mismatch cleavage analysis

Base differences in target DNA were initially detected by a modification of the technique described by Cotton et al. [18,19]. Briefly, PCR was performed on all target DNAs, plus the negative control 'probe' sample. Approx. 100 ng of purified, amplified DNA from the 'probe' individual, was then end-labelled with 10 pM [γ-^{32}P] ATP (< 5000 Ci/mmol, Amersham, U.K.) using T4 polynucleotide kinase (BRL, Paisley, U.K.) and the labelled products were purified by centrifugation through a Sephadex G-25 (Pharmacia, Uppsala, Sweden) column in a vol. of 100 μl 10 mM Tris (pH 7.6)/1 mM Na_2 EDTA (TE) buffer.

Heteroduplex DNA was formed by mixing 10 ng of labelled 'probe' to 100 ng purified, amplified target DNA in 50 μl of 0.3 M NaCl/0.1 M Tris (pH 8.0), boiling for 5 min then incubating for 1 h at 65°C. Duplex DNA, was ethanol precipitated overnight at −20°C with the addition of 40 μg mussel glycogen

(Boeringer-Mannheim, U.K.) as a carrier. The resulting pellet was washed once with ice-cold 70% ethanol and dried briefly before resuspension in 16 μl TE buffer. 7 μl of the heteroduplex solution was made up to 20 μl by the addition of 4 mM hydroxylamine hydrochloride solution (Aldrich, U.K.) (titrated to pH 6.0 with diethylamine (Aldrich)). This was incubated at 37°C for 2 h. The remaining heteroduplex DNA was added to 16 μl of a buffer consisting of 5 mM Tris (pH 8.0), 0.5 mM EDTA, 3% pyridine and 0.025% osmium tetroxide (Aldrich, U.K.) and incubated at 37°C for 2 h. DNA from both reactions was ethanol precipitated overnight, and after washing was resuspended in 50 μl 1 M piperidine (Aldrich) and incubated at 90°C for 45 min. After a final ethanol precipitation, samples were resuspended in 2 μl stop solution (as for sequencing) and run on denaturing polyacrylamide gels in the same manner.

Results

All patients investigated in this study had been shown to be moderately to severely hypercholesterolaemic on several occasions, even after dietary advice. Some of the biochemical features of these individuals are shown in Table I. Seven of the eight (patients 1–7) have had LDL FCR determined by in vivo turnover studies with autologous LDL [20–22]. Three of the eight (patients 6, 7 and 8) have demonstrated significantly reduced growth propagation in the U937 cell assay with duplicate plasma samples taken several months apart. Thus LDL from all patients in this study displays evidence for reduced binding to the normal LDL-receptor in one or both of the assays. Apo B RFLP genotype of the patients was determined by Southern blotting and the results are shown in Table I.

The entire 1544 bp gene region studied in these patients was amplified using two overlapping sets of PCR primers, designated: A1/A2 and B3/B4. (Fig. 1). Samples 1–5 and 8 were initially examined by PCR amplification using oligonucleotides A1/A2, followed by cloning and sequencing. Ten clones from each subject were sequenced across this 800 bp stretch using three primers: the M13 universal primer, Sa and Sb (Fig. 1). Although many PCR errors were found [17], no base change was seen to occur in more than one clone from any patient or in more than one patient. Fig. 2 shows an example of such an experiment from patient 8, where compared to all the others, clone 3 has two bases absent from the 'G' track and two corresponding additional bases in the 'A' track. Since a heterozygous base change would be expected to be seen in approx. half the clones derived from one patient, we concluded that none of the patients studied had a mutation in this 800 bp region.

All patients 1 to 8 were then screened by chemical mismatch cleavage in the gene region bounded by

Table I

Patient characteristics

	Serum total cholesterol (mmol/l)	Serum LDL cholesterol (mmol/l)	Serum triacylglycerols (mmol/l)	FCR [d] (pool/day)	U937 cell [f] growth (%)	Genotype		
						*Xba*I	*Msp*I	*Eco*Ri
1 [a]	8.33 ±0.22	5.46 ±0.07	2.23 ±0.40	0.218	–	22	11	11
2 [a]	8.56 ±0.47	6.2 ±0.60	2.76 ±0.84	0.197	–	22	12	11
3 [a]	6.36 ±0.19.	3.76 ±0.22	0.82 ±0.05	0.182	–	22	11	11
4 [a]	8.00 ±0.38	5.62 ±0.39	3.11 ±0.33	0.265	–	22	11	11
5 [a]	7.03 ±0.20	4.88 ±0.24	1.68 ±0.19	0.217	–	22	11	11
6	9.32	5.79	3.91	0.258 [e]	43/17/35	22	11	11
7	7.61	5.00	3.27	0.302 [e]	42/68/74	22	11	11
8	7.09	4.86	1.01	–	49/43	12	11	11
9 [a,b]	7.08 ±0.15	5.28 ±0.36	1.88 ±0.65	0.297	–	11	11	11
10 [c]	8.0	6.2	0.80	–	22/25	12	12	12

[a] For patients 1, 2, 3, 4, 5 and 9 the lipid levels shown are the mean of three determinations.

[b] Patient 9 is the individual of genotype X-X- at the *Xba*I RFLP site, used as the negative control in all the experiments and the 'probe' in the chemical mismatch analysis.

[c] Patient 10 is a carrier of the apo B_{3500} mutation, used as a positive control in the chemical mismatch analysis.

[d] FCR in normal controls = 0.35 ± 0.06 pools/day [12].

[e] FCR determined by urine plasma ratio method [21,22] and not directly comparable to the results from patients 1 to 5.

[f] Results of repeat determinations on samples taken one month apart are given separated by /.

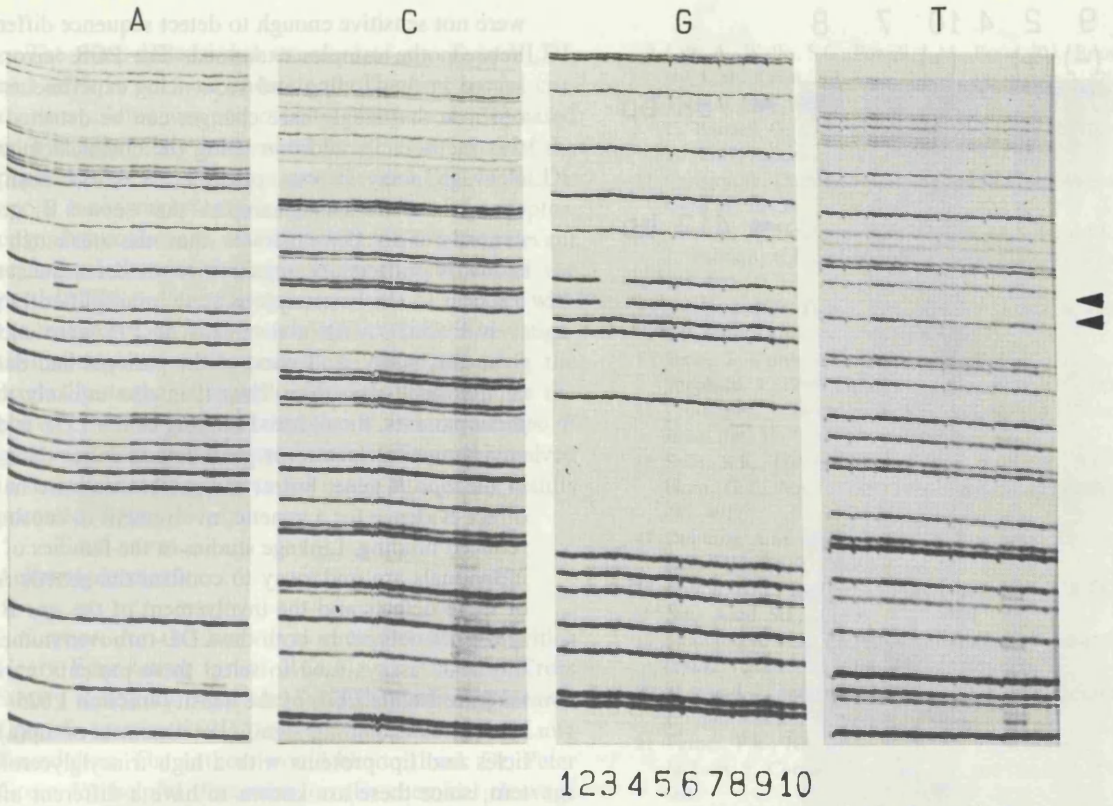


Fig. 2. Example of cloning and sequencing: ten clones derived from patient 8, sequenced by priming with oligonucleotide **Sa**. The sequence of nucleotides 9981 to 10100 are shown. Two G to A misincorporation errors generated during the PCR reaction can be seen in clone 3 – indicated by arrows. The order of bases is given along the top and the clone number along the bottom of the G tracks, the order is the same in each set of tracks.

amplimers A1/A2. No bands were seen on the gel that were not present in the negative control sample, confirming our findings of no sequence differences in this region in these patients (results not shown).

A search was made for sequence differences between bases 10355 and 11143 by PCR amplification using oligonucleotides B3/B4 followed by chemical mismatch cleavage. Amplified DNA from a patient previously demonstrated to carry the G to A base change creating the apo B_{3500} mutation [6] was used as a positive control for a base change in the chemical mismatch cleavage assay (individual 10, Table I). In addition, both this patient and patient 2 are heterozygous for the *Msp*I RFLP in this region of the gene [23] (Fig. 1), and the G to A base change creating this also served as a positive control for the mismatch assay. Hydroxylamine modifies mismatched deoxycytidine in the 'probe' DNA, which corresponds to a nucleotide other than guanosine in the appropriate position, on the opposite strand in the test DNA. A typical result is shown in Fig. 3; both positive controls show the expected cleavage products, with the DNA from patient 2 (heterozygous for the *Msp*I RFLP) producing an 86 bp band with hydroxylamine, and that from the patient with the apo B_{3500} mutation generating fragments of 412 and 86 bp when modified with hydroxylamine. No other bands corresponding to any sequence differences were seen in any of the samples.

Discussion

We set out to identify common sequence differences in the *apo B* gene that are in linkage disequilibrium with the apo B *Xba*I RFLP, and that cause alterations in plasma lipid levels in healthy individuals [8–10,24,25] by comparing the sequence differences between the X- alleles and the X + alleles. We expect that this sequence difference will not have a large impact on the serum cholesterol levels of any one individual, and since it is common, will be present in at least one of the 18 apo B alleles whose sequence was investigated in this study. All the patients recruited to this study demonstrated reduced binding of their LDL to a normal LDL-receptor in one or other of the assays used [12,13,15]. However, in none of the alleles examined did we detect sequence differences in the region of the *apo B* gene coding for the putative receptor binding domain, amino acids 3130–3630. We therefore conclude that the sequence variations causing defective binding in these patients are outside this region of the *apo B* gene.

Maeda et al. [26] have demonstrated similar results in swine, where an allele of apo B, termed Lpb5.1, is

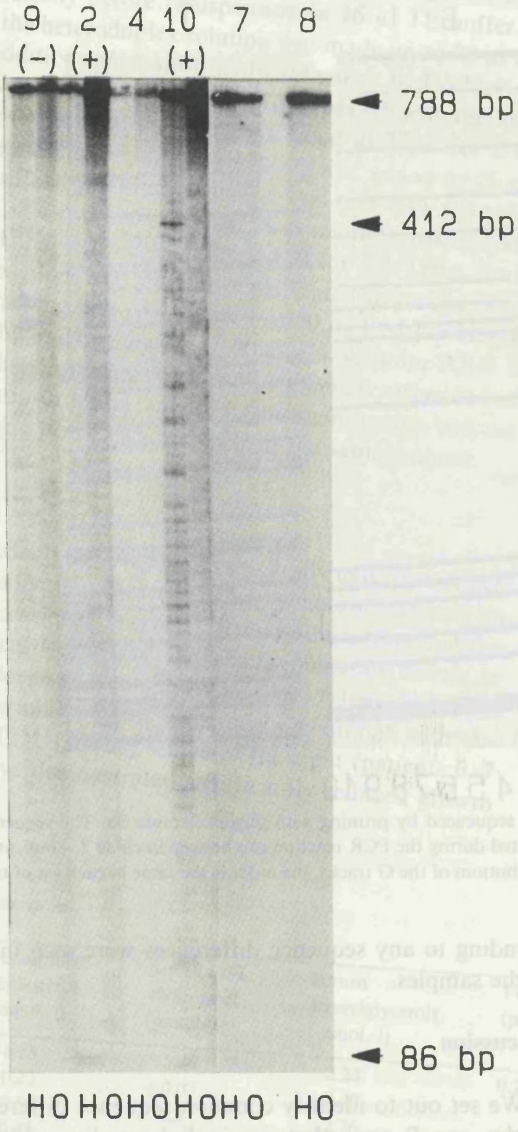


Fig. 3. Example of chemical mismatch cleavage analysis of samples amplified with amplimers B3/B4. Patients' code numbers are given along the top. There are two tracks per patient: the hydroxylamine reaction (**H**) and the osmium tetroxide reaction (**O**). The uncleaved, full-length PCR gives a band of 788 bp. DNA from patient 9 was used as the 'probe' for all the other samples and as the negative control (−). DNA from patients 2 and 10 (+) shows an 86 bp band corresponding to the base change creating the *Msp*I RFLP and DNA from patient 10 also gives a 412 bp band arising from the base change creating the apo B_{3500} mutation.

associated with reduced clearance of LDL, hyperlipidaemia and atherogenesis. These workers sequenced the region of the gene coding for amino acids 3133–3497 (the putative receptor binding domain in swine apo B) and found no differences between this allele and three other alleles which were not associated with the same phenotype. This implies that sequences outside this region must also affect interaction of LDL with the LDL-receptor in pigs.

It is unlikely that the techniques used in this study were not sensitive enough to detect sequence differe between the samples examined. The PCR 'errors' served in the cloning and sequencing experiments d onstrate that single base changes can be detected u this method. In addition using the chemical mism cleavage assay it was possible to detect both changes in the DNA samples that served as pos controls. This demonstrates that the mismatch t nique is sufficiently sensitive to detect sequence ferences in the heterozygous state in amplified DN

Individuals with a diagnosis of FH were exclu from this study, and none of the patients had diab or thyroid disfunction. Thus it is also unlikely tha these patients, the reduced binding of the LDL is du environmental factors or gene defects other than t in the *apo B* gene, but at the present time we hav direct evidence for a genetic involvement in causing reduced binding. Linkage studies in the families of t individuals are underway to confirm the genetic na of these defects and the involvement of the *apo B* in the phenotype. In both the LDL-turnover studies the U937 assays used to select these patients, care taken to isolate LDL of the density fraction 1.025–1 kg/l, thus excluding significant amounts of Lp(a) ticles and lipoproteins with a high triacylglycerol tent, since these are known to have a different aff for the normal LDL-receptor in vitro [27].

The original receptor binding domain (amino a 3147–3381) was proposed on the basis of a numbe different criteria [1,2], but there are several reason believe that amino acid substitutions outside this c nally proposed region can affect binding of LDL t LDL-receptor. The strongest evidence comes fror mutation at amino acid 3500 that causes FDB glutamine for arginine substitution that creates thi directly affect the interaction between the ligan main of apo B and the LDL-receptor or it may ha indirect affect on binding by altering the conforn of the protein in this region. Evidence tha carboxyl-terminal end of the protein is invol modulating binding to the LDL-receptor has com the study of a patient with an apo B protein truncated at residue 4041, and thus is only 89% full length protein (designated B89) [28]. This has hypobetalipoproteinaemia due to the rapid cl of his apo B89 containing LDL which has ir affinity to the LDL-receptor on normal fibrobla suggesting that residues in the deleted COOH- portion of apo B may also have a role in the in of apo B with its receptor. A third line of evid come from the work of Chapman and his c using monoclonal antibodies raised to short pe apo B, which demonstrated that an antibody r peptide consisting of amino acids 4007 to inhibit binding and degradation of LDL b LDL-receptors [30].

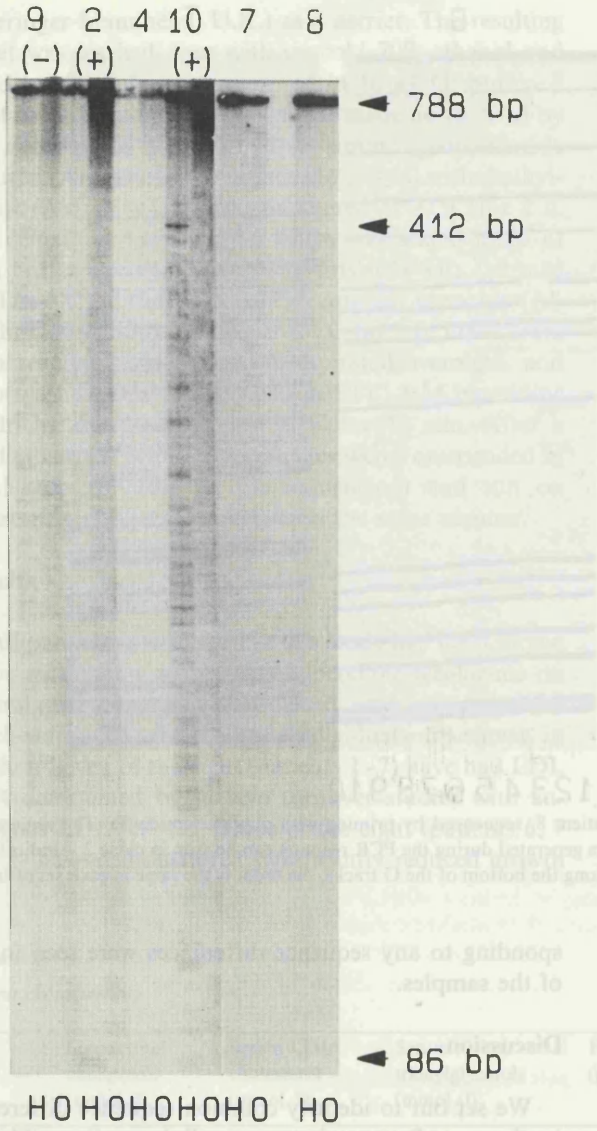


Fig. 3. Example of chemical mismatch cleavage analysis of samples amplified with amplimers B3/B4. Patients' code numbers are given along the top. There are two tracks per patient: the hydroxylamine reaction (**H**) and the osmium tetroxide reaction (**O**). The uncleaved, full-length PCR gives a band of 788 bp. DNA from patient 9 was used as the 'probe' for all the other samples and as the negative control (−). DNA from patients 2 and 10 (+) shows an 86 bp band corresponding to the base change creating the *Msp*I RFLP and DNA from patient 10 also gives a 412 bp band arising from the base change creating the apo B_{3500} mutation.

associated with reduced clearance of LDL, hyperlipidaemia and atherogenesis. These workers sequenced the region of the gene coding for amino acids 3133–3497 (the putative receptor binding domain in swine apo B) and found no differences between this allele and three other alleles which were not associated with the same phenotype. This implies that sequences outside this region must also affect interaction of LDL with the LDL-receptor in pigs.

It is unlikely that the techniques used in this study were not sensitive enough to detect sequence differences between the samples examined. The PCR 'errors' observed in the cloning and sequencing experiments demonstrate that single base changes can be detected using this method. In addition using the chemical mismatch cleavage assay it was possible to detect both base changes in the DNA samples that served as positive controls. This demonstrates that the mismatch technique is sufficiently sensitive to detect sequence differences in the heterozygous state in amplified DNA.

Individuals with a diagnosis of FH were excluded from this study, and none of the patients had diabetes or thyroid disfunction. Thus it is also unlikely that in these patients, the reduced binding of the LDL is due to environmental factors or gene defects other than those in the *apo B* gene, but at the present time we have no direct evidence for a genetic involvement in causing this reduced binding. Linkage studies in the families of these individuals are underway to confirm the genetic nature of these defects and the involvement of the *apo B* gene in the phenotype. In both the LDL-turnover studies and the U937 assays used to select these patients, care was taken to isolate LDL of the density fraction 1.025–1.050 kg/l, thus excluding significant amounts of Lp(a) particles and lipoproteins with a high triacylglycerol content, since these are known to have a different affinity for the normal LDL-receptor in vitro [27].

The original receptor binding domain (amino acids 3147–3381) was proposed on the basis of a number of different criteria [1,2], but there are several reasons to believe that amino acid substitutions outside this originally proposed region can affect binding of LDL to the LDL-receptor. The strongest evidence comes from the mutation at amino acid 3500 that causes FDB. The glutamine for arginine substitution that creates this may directly affect the interaction between the ligand domain of apo B and the LDL-receptor or it may have an indirect affect on binding by altering the conformation of the protein in this region. Evidence that the carboxyl-terminal end of the protein is involved in modulating binding to the LDL-receptor has come from the study of a patient with an apo B protein that is truncated at residue 4041, and thus is only 89% of the full length protein (designated B89) [28]. This patient has hypobetalipoproteinaemia due to the rapid clearance of his apo B89 containing LDL which has increased affinity to the LDL-receptor on normal fibroblasts [29], suggesting that residues in the deleted COOH-terminal portion of apo B may also have a role in the interaction of apo B with its receptor. A third line of evidence has come from the work of Chapman and his colleagues using monoclonal antibodies raised to short peptides of apo B, which demonstrated that an antibody raised to a peptide consisting of amino acids 4007 to 4019 will inhibit binding and degradation of LDL by normal LDL-receptors [30].

The three-dimensional structure of apo B in VLDL and LDL is not well understood. The best model currently available was proposed by Yang et al. [31] based on hydrophobicity studies and the determination of the peptides that are trypsin accessible in LDL. In VLDL apo B is not available as a ligand for the LDL-receptor, only after conversion of VLDL to LDL are the relevant regions of the protein exposed on the surface of the particle [32] and are therefore available to interact with the receptor. Our future work will involve the investigation of regions of the gene 3′ to those encoding the originally envisaged receptor binding domain since the above evidence suggests that the 3′ end of the *apo B* gene may encode the other domains of apo B involved in receptor binding, postulated in the light of the results presented here.

Acknowledgements

Support for this work was provided by the British Heart Foundation (grant RG 5), the Charing Cross Sunley Research Trust, the Swedish Medical Research Council (grant 08691) and the Swedish Heart-Lung Foundation. The authors would like to thank Dr. Peter Green for helpful discussion on the chemical cleavage mismatch analysis and Ms Elaine Osman for help in preparing the manuscript.

References

1 Yang, C-Y., Chen, S-H., Sparrow, J.T., Gianturco, S-H., Bradley, W.A., Tanimuro, M., Li, W-H., Sparrow, D.A., DeLoof, H., Rosseneu, M., Proot, L., Brasseur, B., Ruysschaert, J.M., Lee, F-S., Gu, Z-W., Gotto, A.M. and Chan, L. (1986) Nature 323, 738–742.
2 Knott, T.J., Pease, R.J., Powell, L.M., Wallis, S.C., Rall, S.C., Innerarity, T.L., Blackhart, B., Taylor, W.H., Marcel, Y., Milne, R. et al. (1986) Nature 323, 734–738.
3 Innerarity, L.H., Weisgrabber, K.H., Arnold, K.S., Mahley, R.W., Krauss, R.M., Vega, G.L. and Grundy, S.M. (1987) Proc. Natl. Acad. Sci. USA 84, 6919.
4 Soria, L.F., Ludwig, E.H., Clarke, H.R., Vega, G.L., Grundy, S.M. and McCarthy, B.J. (1989) Proc. Natl. Acad. Sci. USA 86, 587–591.
5 Vega, G.L. and Grundy, S.M. (1986) J. Clin. Invest. 78, 1410–1414.
6 Tybærg-Hansen, A., Gallagher J., Vincent., Houlston R. H., Talmud P., Dunning, A.M., Seed, M., Hamsten, A., Humphries, S.E. and Myant, N.B. (1990) Atherosclerosis 80, 235–242.
7 Schuster, H., Rauh, G., Kormann, B., Hepp, T., Humphries, S., Keller, C., Wolfram, G. and Zöllner, N. (1990) Arteriosclerosis 10, 577–581.
8 Berg, K. (1986) Clin. Genet. 30, 515–520.
9 Law, A., Wallis, S.C., Powell, L.M., Pease, R.J., Brunt, H., Priestley, L.M., Knott, T.J. et al. (1986) Lancet June 7, 1301–1303.
10 Talmud, P.J., Barni, N., Kessling, A.M., Carlsson, P., Darnfors, C., Bjursell, G., Galton, D., Wynn, V., Kirk, H., Hayden, M.R. et al. (1987) Atherosclerosis 67, 81–89.
11 Carlsson, P. Darnfors, C., Olofsson, S.O. and Bjursell, G., (1987) Gene 49, 29–34.
12 Demant, T., Houlston, R.S., Caslake, M.J., Series, J.J., Shepherd, J., Packard, C.J. and Humphries, S.E. (1988) J. Clin. Invest. 82, 797–802.
13 Houlston, R.S., Turner, P.R., Revill, J., Lewis, B. and Humphries, S.E. (1988) Atherosclerosis 71, 81–85.
14 Series, J., Cameron, I., Caslake, M., Gaffney, D., Packard, C. and Shepherd, J. (1989) Biochim. Biophys. Acta 1003, 183–188.
15 Frostegård, J., Hamsten, A., Gidlund, M. and Nilsson, J. (1990) J. Lipid Res. 31, 1, 33–34.
16 Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higushi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science 239, 487–491.
17 Dunning, A.M., Talmud, P. and Humphries, S.E. (1988) Nucleic. Acids. Res. 16, 21, 10393.
18 Cotton, R.G., Rodrigues, N.R. and Campbell, R.D. (1988) Proc. Natl. Acad. Sci., USA 85, 4397–4401.
19 Montandon, A.J., Green, P.M., Gianelli, F. and Bentley, D.R. (1989) Nucleic. Acids Res. 17, 9, 3347–3358.
20 Shepherd, J., Bicker, J.S., Lorimer, A.R. and Packard, C.J. (1979) J. Lipid Res. 20, 999–1006.
21 Turner, P.R., Konarska, R., Revill, J., Masana, L., La Ville, A., Jackson, P., Cortese, C., Swan, A.V. and Lewis, B. (1984) Lancet ii 663.
22 International Collaborative Study Group (1986) Lancet ii 991.
23 Huang, L.H., De Graaf, J. and Breslow, J.L. (1988) J. Lipid. Res. 239, 63–67.
24 Alto-Setälä, Gylling, H., Helve, E., et al. (1989) Hum. Genet. 82, 305–307.
25 Paulweber, P., Friedl, W., Krempler, S., Humphries, S.E. and Sandhofer, F. (1990) Arteriosclerosis 10, 1, 17–24.
26 Maeda, N., Ebert, D.L., Doers, T.M., Newman, M., Hasler-Rapacz, J., Attie, A.D., Rapacz, J. and Smithies, O. (1988) Gene 70, 213–229.
27 Kleinman, Y., Eisenburg, D., Oschry, Y., Gavish, D., Stein, O. and Stein, Y. (1985) J. Clin. Invest. 75, 1796–1803.
28 Talmud, P.J., King-Underwood, L., Krul, E., Dunning, A., Schonfeld, G. and Humphries S.E. (1989). J. Lipid Res. 30, 1773–1780.
29 Krul, E.S., Kinoshita, M., Talmud, P., Humphries, S.E., Turner, S., Goldberg, A.C., Cook, K., Boerwinkle, E. and Schonfeld, G. (1989) Arteriosclerosis 9, 856–868.
30 Forgez, P., Rouis, H., Guo, H-C., Nigon, F. and Chapman, M.J. (1988) Circulation 78, 4: II–286: 1141, abstact.
31 Yang, C-Y., Gu, Z-W., Weng, S-A., Kim, T.W., Chen, S-H., Pownall, H.J., Sharp, P.M., Liu, S-W., Li, W-H., Gotto, A.M. and Chan, L. (1989) Arteriosclerosis 9, 1, 96–108.
32 Bradley, W.A., Hwang, S.L., Karlin, J.B., Lin, A.H., Prasad, S.C., Gotto, A.M. and Gianturco, S.H. (1984) J. Biol. Chem. 259, 14728–35.

Am. J. Hum. Genet. 50:208–221, 1992

Two Amino Acid Substitutions in Apolipoprotein B Are in Complete Allelic Association with the Antigen Group (x/y) Polymorphism: Evidence for Little Recombination in the 3′ End of the Human Gene

Alison M. Dunning,*,[1] Helmut-H. Renges,*,[1] Chun-Fang Xu,*,[1] Rachel Peacock,*,[1] Robert Brasseur,‡ Gerald Laxer,† Matti J. Tikkanen,§ Réné Bütler,‖ N. Saha,# Anders Hamsten,** Maryvonne Rosseneu,†† Philippa Talmud,*,[1] and Steve E. Humphries*,[1]

*Charing Cross Sunley Research Centre; and †Department of Crystallography, Bickbeck College, London; ‡Laboratory of Macromolecules at Interfaces, Free University of Brussels, Brussels; §Third Department of Medicine, University of Helsinki, Helsinki; ‖Central Laboratory Blood Transfusion Service, Swiss Red Cross, Berne; #Department of Pediatrics, National University of Singapore, Singapore; **King Gustav V Research Institute, Karolinska Institute and Department of Internal Medicine, Karolinska Hospital, Stockholm; and ††Department of Clinical Biochemistry, A.Z. St-Jan, Brugge, Belgium

Summary

We report the identification of an A-to-G base change, in exon 29 of the apolipoprotein B (apo B) gene, that results in the substitution of serine for asparagine at residue 4311 of mature apo B100. In a recent publication, Huang et al. have reported a C-to-T base change in exon 26 that causes the substitution of leucine for proline at residue 2712 of apo B. We have found complete linkage disequilibrium between the alleles at both these sites and an immunochemical polymorphism of LDL designated antigen group (x/y) (Ag(x/y)) in a sample of 118 Finnish individuals. This implies that either one of these substitutions—or both of them combined—could be the molecular basis of the Ag(x/y) antigenic determinants, with the allele encoding $serine_{4311}$ plus $leucine_{2712}$ representing the Ag(x) epitope, and that encoding $asparagine_{4311}$ plus $proline_{2712}$ the Ag(y) epitope. In a sample of 90 healthy Swedish individuals the Leu_{2712}/Ser_{4311} allele is associated both with reduced serum levels of LDL-cholesterol and apo B and with raised levels of HDL. However, these differences are of smaller effect than those associated with the *Xba*I RFLP of the apo B gene in this sample. We have also genotyped 523 individuals from European, Asian, Chinese, and Afro-Caribbean populations and have found complete association between the sites encoding residues 2712 and 4311 in all of these samples, although there are large allele frequency differences between these populations. In addition, there is strong linkage disequilibrium with allelic association between the alleles of these sites and those of the *Xba*I RFLP in all the populations examined. Taken together, these data suggest that, since the divergence of the major ethnic groups, there has been little or no recombination in the 3′ end of the human apo B gene.

Received February 28, 1991; final revision received August 19, 1991.

Address for correspondence and reprints: Alison M. Dunning, Centre for the Genetics of Cardiovascular Disorders, Department of Medicine, The Rayne Institute, University Street, London WC1E 6JJ, United Kingdom.

1. Present address: Centre for Genetics of Cardiovascular Disorders, Department of Medicine, The Rayne Institute, University Street, London WC1E 6JJ, United Kingdom.

0002-9297/92/5001-0023$02.00

Introduction

Apolipoprotein B100 (apo B) is a 550-kD protein secreted by the liver as a constituent of VLDL. During the metabolism of VLDL, other apoproteins and triglycerides are removed, leaving apo B as the sole protein component of LDL. As such, it serves both to maintain the integrity of the particle (Yang et al. 1986, 1989), and, as the ligand for the LDL receptor, it mediates the clearance of LDL-cholesterol from the

plasma. Since elevated plasma levels of LDL-cholesterol are one of the recognized risk factors for the development of coronary artery disease (CAD) (Whayne et al. 1981; Durrington et al. 1986), the identification of the factors that determine plasma LDL-cholesterol levels are of major public health importance. Twin studies and path analysis have demonstrated that genetic variation has a significant impact on both plasma levels of apo B and LDL-cholesterol, with a heritability of .5–.6 (Hamsten et al. 1986*a*; Berg 1987). Family studies using complex segregation analysis have found evidence for the plasma levels of apo B being determined by a major gene with contributions from environmental factors and genes of small or intermediate effect (Hasstedt et al. 1987; Pairitz et al. 1988).

Protein polymorphisms of apo B were first noted as the antigen-group (Ag) system of variants, detected using antisera from multiply transfused patients (Allison and Blumberg 1961). There are five reported pairs of epitopes, each pair representing alleles of the apo B gene; these are Ag(a_1/d), Ag(c/g), Ag(h/i), Ag(t/z), and Ag(x/y) (reviewed in Breguet et al. 1990). Berg et al. (1976), using the combined data from 10 different populations, reported that one of the epitopes of this system, the Ag(x−) (equivalent to Ag(y)), is associated with raised serum cholesterol and triglyceride levels. Since the cloning of the human apo B gene (Blackhart et al. 1986), nucleotide substitutions have been reported as candidates for the molecular bases of all the Ag epitopes (Ma et al. 1987, 1989; Dunning et al. 1988; Wang et al. 1988; Young and Hubl 1989; Xu et al. 1989; Wu et al. 1991).

An RFLP of the apo B gene, detected by the enzyme *Xba*I, has been documented by numerous groups to be associated with differences in serum cholesterol and apo B levels, with the presence of the cutting site (the X+ allele) being associated with raised levels (Berg et al. 1986; Law et al. 1986; Talmud et al. 1987; Aalto-Setälä et al. 1989; Paulweber et al. 1990). In addition, the same allele has been found to be associated with reduced fractional catabolic rate (FCR) of LDL (Demant et al. 1988; Houlston et al. 1988). However, the *Xba*I RFLP cannot be the direct cause of the effects seen, since the base change that creates the variant *Xba*I site is in the third base (wobble position) of codon 2488 and so does not alter the threonine residue at this position (Carlsson et al. 1986). The alleles of this RFLP must be in linkage disequilibrium with another base change that does have a functional effect leading to the reported differences in lipid parameters. The present study set out to identify common base changes, in linkage disequilibrium with the alleles of the *Xba*I site, that alter apo B structure and thus may be the cause of the associations seen.

Subjects and Methods

1. *Subjects*

a. British.—These subjects were two groups of unrelated hyperlipidemic individuals on whom FCR of LDL studies have been performed. FCR results and the apo B genotyping of the 17 individuals from Glasgow and of 22 individuals from Sheffield have been described elsewhere (Demant et al. 1988; Houlston et al. 1988).

b. North Karelia, Finnish.—This group comprised 89 unrelated men and women, with no history of hyperlipidemia, from a rural community in North Karelia, Finland. This population has been very stable for several hundred years, with no influence of Swedish-speaking individuals. They have been described in detail elsewhere (Ehnholm et al. 1982, 1984; Kuusi et al. 1985). Ag phenotyping and apo B RFLP genotyping have also been described elsewhere (Tikkanen et al. 1989; Xu et al. 1989).

c. Other Finnish.—A second group of Ag phenotyped, unrelated Finnish individuals was also examined. Phenotyping and apo B genotyping have been described by Dunning et al. (1988).

d. Swedish.—This group comprised 95 randomly selected, healthy male residents of Stockholm county. Data on lipid, lipoprotein, apoprotein, hemostatic, and metabolic variables have been presented elsewhere (Hamsten et al. 1986*b*, 1987). Individuals with serum cholesterol levels >9.5 mM and/or triglyceride levels >3.0 mM were excluded from the lipid and lipoprotein analyses. Age and body mass index (BMI) accounted, respectively, for 11.5% and 7.6% of the variance in serum HDL-cholesterol and triglyceride levels in this group.

e. Chinese.—This group comprised 83 healthy Chinese males, mean age 36 years, living in Singapore. A detailed description of this sample is presented elsewhere (Saha et al., submitted).

f. Afro-Caribbeans.—This group comprised 47 individuals aged 45–74 years (mean age 55 years), with at least three grandparents of Afro-Caribbean origin, drawn from family-practitioner registers in two Northwest London health centers. This sample has been described in detail elsewhere (Miller 1989).

g. Asians.—This group was a random subset of 152 men drawn from a sample of 1,433 south Asian men aged 40–65 years and living in the west of London who have been investigated for CHD risk factors. The original sample was assembled from four factories and from lists of general practitioners in west London. Of these men, 89% were Pujabi Sikhs, and the rest were evenly divided between Gujarati and Punjabi Hindus (McKeigue et al. 1991).

Apes.—DNA from four unrelated chimpanzees, four unrelated gorillas, and two orangutans was also investigated.

2. *DNA Isolation and Southern Blot Analysis*

These procedures were performed using standard techniques described elsewhere (Talmud et al. 1987).

3. *PCR*

Regions of the apo B gene were amplified by PCR (Saiki et al. 1988). The sequences of the oligonucleotide amplimers used, as well as their arrangements, are shown in figures 1 and 3. Oligonucleotides were obtained from Severn Biotech Ltd., Kidderminster, UK. The PCR reactions were performed in an automated thermal cycler (Cambio, Cambridge) by using *Thermus aquaticus* DNA polymerase (Perkin Elmer Cetus, Norwalk, CT) in the manufacturer's recommended buffer. After both initial denaturing of the DNA at 95°C for 5 min and annealing of the primers (55°C, 3 min), the program used was 50 cycles of 72°C for 3 min, 95°C for 1 min, and 55°C for 1 min. DNA generated by the PCR reactions was purified by Geneclean (Bio101, La Jolla, CA) according to the manufacturer's instructions and was eluted into TE buffer to give a DNA concentration of 50–100 ng/μl.

4. *Chemical Mismatch Cleavage Analysis*

Base differences in target DNA were initially detected by a modification of the Cotton et al. (1988) technique, described by Montandon et al. (1989). The exact procedure followed is described by Dunning et al. (1991).

5. *Direct Sequencing*

PCR was performed, and the products were purified as described above. Approximately 100 ng of purified product was annealed to a 30-fold molar excess of oligonucleotide primer, internal to the PCR primers (Severn Biotech, Kinderminster, UK) and was sequenced as described by King-Underwood et al. (1991).

6. *Blotting and Hybridization to Allele-specific Oligonucleotides (ASOs)*

PCRs were performed as described above but were not purified. Oligonucleotides (Severn Biotech) used as PCR primers, along with ASOs, are listed in figures 1 and 3. ASOs were labeled, at the 5′ end, with T4 polynucleotide kinase (BRL, Paisley, UK) and with [γ-^{32}P] ATP (Amersham, Amersham, UK) to a specific activity of approximately 0.1 μCi/pmol (Maniatis et al. 1982). Initially, slot blotting was performed by applying one-fiftieth of each PCR reaction to nylon filters (Hybond-N; Amersham, Amersham, UK) by using a slot-blotting manifold (Schleicher & Schuell, Dossel, Germany), but later one-tenth of the PCR was run on 1.2% agarose gels, was denatured in 1.5 M NaCl/0.5 M NaOH, and then was double-blotted between two pieces of nylon membrane (Biodyne A, Pall Processing Ltd, Portsmouth, UK), by using the denaturing solution as the transfer buffer. DNA was bound to the filters by baking at 80°C for 2 h. Filters were hybridized for 3 h in 5 × SSPE/0.5% SDS/5 × Denhardt's solution, with one or other of the ASOs, at 33°C and then was washed in 5 × SSPE/0.1% SDS for 10 min at the melting temperature of the ASO. Autoradiography was for 2–16 h.

7. *Computer Estimation of Effects of Amino Acid Substitutions*

The protein secondary structure was modeled using the Predict suite of 10 secondary structure prediction programs (Department of Biophysics, University of Leeds). Secondary structure predicted by this program is approximately 60% reliable. Segments of amino acid sequence extending 250 residues either side of the polymorphic sites were analyzed. Hydrophobicity profiles were generated for the same segments by using the Pc-Prot+ (Protein Analysis) Program (1990 version) and by calculating the mean hydrophobic moment (μ_H) and mean hydrophobicity (H_i), using a seven-residue window, according to the procedure of Eisenberg (1984). μ_H was then plotted against H_i for segments extending 15 amino acid residues either side of the relevant substitution sites (Rosseneu et al. 1990).

8. *Statistical Methods*

The gene-counting method was used for the estimation of gene frequencies. χ^2 Analysis was used for the estimation of Hardy-Weinberg equilibrium. Between-sample differences in apo B genotype distribu-

tion were determined by χ^2 analysis. Association between genotypes at the polymorphic loci was also estimated by χ^2 analysis. We considered statistical significance to be at the .05 level. The strength of associations was estimated by the correlation coefficient Δ (Chakravarti et al. 1984). For each lipid trait, multiple linear regression was used to adjust for age, gender, and BMI. The skewness of triglyceride distribution was adjusted by $\log_e$ transformation. For each adjusted variable, a one-way analysis of variance was performed to test the null hypothesis that phenotypic variation is not associated with genetic variation at the apo B locus. The percentage of variance ($R^2 \times 100$) explained by apo B genotypes was estimated by nonlinear regression.

Results

The chemical cleavage mismatch analysis technique was used to screen amplified DNA at bases 12326–13379 in exon 29 of the apo B gene from nine hyperlipidemic British patients. Seven of these patients (1–7 in fig. 1) are homozygous for the presence of the apo B *Xba*I cutting site and have been shown to have reduced FCR of LDL when compared with a similar group of patients homozygous for the absence of the *Xba*I cutting site (Demant et al. 1988; Houlston et al. 1988; Dunning et al. 1991). Patient 8 has been demonstrated to have LDL with a reduced affinity for a normal LDL receptor in an in vitro cell-binding assay (Frostegård et al. 1990; Dunning et al. 1991), and he is heterozygous for the *Xba*I RFLP. The sequence of the DNA amplified from these eight patients was compared with that from the probe DNA amplified from a subject homozygous for the absence of the *Xba*I site and with increased FCR of LDL as demonstrated in the study by Demant et al. (1988). The aim of the present experiment was to find a base change common to the X+ alleles—but not present in the X− alleles—that may be the cause of the differences, in both LDL-cholesterol levels and FCR of LDL, seen associated with the presence or absence of the *Xba*I cutting site. The exon 29 region examined by chemical cleavage mismatch analysis was covered in two overlapping PCR fragments (fig. 1, *top*).

In the analysis of the second PCR fragment, a novel 230-bp band was seen in the hydroxylamine tracks of all the samples (fig. 1, *bottom*). However, this band was present at reduced intensity in the probe sample and in the DNA from patient 8, suggesting that these two subjects were heterozygous for the base change while the others were homozygous. Direct sequencing was performed on the amplified DNA from these patients, and it was determined that the base change occurring is an A-to-G at nucleotide 13141 (fig. 2). This results in a serine-for-asparagine substitution at residue 4311 of the mature apo B100 protein.

The presence of this A-to-G base change in two of nine subjects suggested that this was a common variant, and thus a sample of unrelated individuals were genotyped to estimate the frequency of this polymorphism. The North Karelia and Finnish samples combined (together totaling 119 unrelated Finnish subjects) were genotyped for this base change by using ASOs (fig. 3, *top*) These subjects had previously been phenotyped for each of the Ag types (Dunning et al. 1988; Tikkanen et al. 1989; Xu et al. 1989). In this sample, the frequency of Asn_{4311} was found to be .79 (table 1), and the distribution of genotypes was in Hardy-Weinberg equilibrium. In addition, when genotype and phenotype data were compared, complete association was found between and $Asn_{4311} \rightarrow$ Ag(x/y) Ser, with the asparagine-encoding allele corresponding to Ag(y) and the serine-encoding allele to Ag(x).

Huang et al. (1990) have reported another polymorphism, in exon 26 of the apo B gene, which generates a substitution of leucine for proline at residue 2712 of the mature protein. Since this amino acid change may have a functional effect on apo B, we examined this polymorphism in the same sample of Finnish individuals (fig. 3, *bottom*). Complete association was also found between this polymorphism and Ag(x/y), with Ag(y) corresponding to Pro_{2712} and Asn_{4311} and Ag(x) with Leu_{2712} and Ser_{4311}. This unexpected finding indicates that either one or both of these loci may form the epitope detected as Ag(x/y).

Since there was complete allelic association between these three polymorphisms in this sample of Finnish individuals, it was decided to investigate this further by determining the genotypes at these sites in samples from other ethnic groups. These included individuals of Swedish, British, Asian, Chinese, and Afro-Caribbean origin (tables 1 and 2). Complete association was found between the alleles encoding these two sites in all of the individuals investigated (523 total). However, the allele frequency of Pro_{2712}/Asn_{4311}/Ag(y) varied from .89 in the Afro-Caribbean sample to .24 in the Chinese (table 1). Our estimations of the allele frequencies at residues 2712 and 4311 are in agreement with those observed for the Ag(x/y) polymorphism in similar ethnic groups (Breguet et al. 1990). Subsequently, Rapacz et al. (1991) have reported that

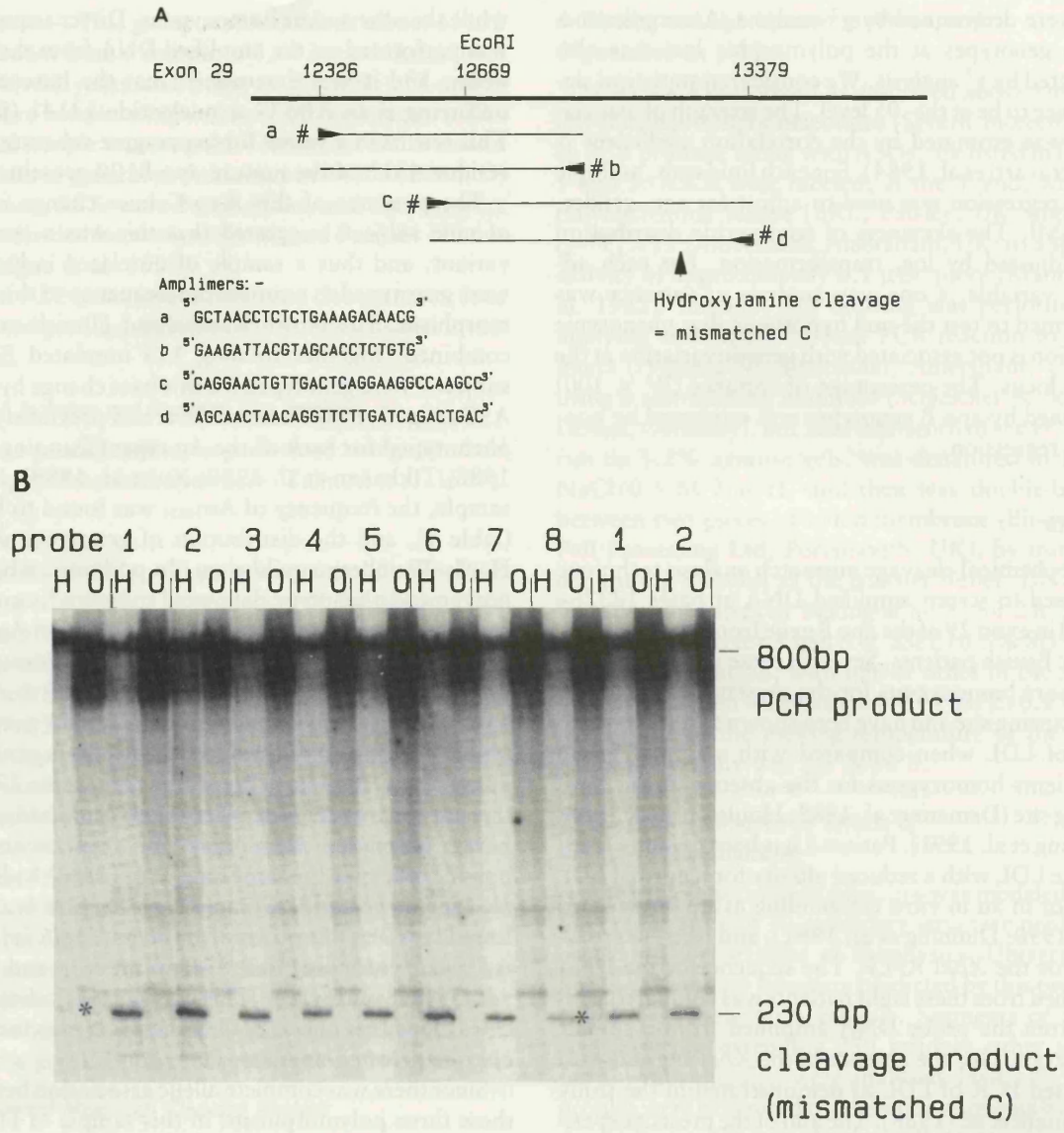


Figure 1 Chemical cleavage mismatch analysis in exon 29 of Apo B Gene. *Top,* Partial map showing gene region covered. Nucleotide numbers are given along the top. Overlapping PCR products used in the analysis are depicted below. Positions of amplimers are represented by arrows, and their sequences are shown as a, b, c, and d. Positions of 5′ end-labeling with ^{32}P are denoted by pound symbols (#). *Bottom,* Autoradiograph showing results of analysis. DNA samples 1–8 and probe (the "probe" DNA sample was hybridized against itself as a negative control) are presented (samples 1 and 2 are shown twice). In each case the hydroxylamine reaction track (H) is on the left, and the osmium tetroxide reaction (O) is on the right. The 800-bp uncleaved PCR product is visible at the top of the photograph, with the 230-bp hydroxylamine cleavage product beneath. The samples yielding reduced intensity bands (i.e., heterozygous for the base change) are marked by asterisks (*).

the Ag(x/y) protein polymorphism was not apparent in 20 chimpanzees and eight gorillas that they investigated, all of which were homozygous for the Ag(y) epitope. Using the identical system that we developed for the human apo B gene, we have genotyped the DNA from four chimpanzees, four gorillas, and two orangutans. Each of these animals was homozygous for proline and asparagine at the apo B residues corre-

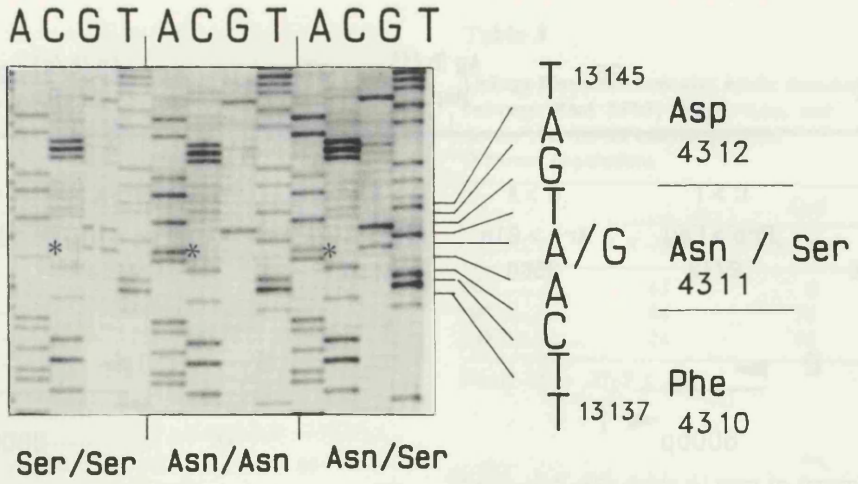


Figure 2 Direct sequencing of DNA amplified across mismatch. Nucleotides 13137–13145 (encoding amino acids 4310–4312) are shown for four DNA samples. The first four tracks show sequenced DNA from an individual homozygous for G at nucleotide 13141 encoding Ser. The second four tracks are from an individual homozygous for A encoding Asn, and the final tracks show DNA from an individual heterozygous for the base difference.

sponding to human residues 2712 and 4311, respectively; this would be compatible with these animals also being homozygous for Ag(y).

In order to try to elucidate what, if any, functional effect the Pro_{2712}→Leu and Asn_{4311}→Ser substitution may have, an attempt was made to determine the secondary structures of the polymorphisms of apo B by using the amino acid sequence data of the relevant portions of the protein. This was done by using both the output of the Predict secondary-structure prediction program and hydrophobicity profiles. The Pro_{2712}→Leu substitution appears to be in a very short helical section within a hydrophobic region of the protein generally comprising β-sheet and β-turn (fig. 4), and it is predicted that the presence of proline at this site would disturb the helical structure but would not significantly affect the hydrophobicity of the region. The Asn_{4311}→Ser change appears to occur in a short joining region between two long amphipathic helices (fig. 4), indicating that it is likely to be exposed on the surface of the LDL particle; however, it would not seem to have a major effect on the secondary structure.

In order to determine whether the polymorphisms at residues 2712 and 4311 are the cause of the differences in lipid and lipoprotein levels seen in association with the *Xba*I RFLP of apo B, we investigated the association between these loci and the *Xba*I site in the Finnish, Swedish, British, Asian, and Chinese samples and found strong linkage disequilibrium ($\Delta = .57$, $P < .001$). Moreover, all the unambiguously determined $Ag(x)/Leu_{2712}/Ser_{4311}$ alleles are associated with the X− allele, although the $Ag(y)/Pro_{2712}/Asn_{4311}$ al-

Table 1

Genotype Distribution and Relative Allele Frequencies in Population Samples

Population (N)	Pro_{2712}→Leu			Asn_{4311}→Ser			Frequency of Pro_{2712}/Asn_{4311}
	PP	PL	LL	NN	NS	SS	
Finnish (118)	65	44	9	65	44	9	.74
British (37)	21	13	3	21	13	3	.74
Swedish Controls (186)	58	23	5	58	23	5	.81
Asians (152)	44	70	38	44	70	38	.52
Chinese (82)	7	34	41	7	34	41	.29
Afro-Caribbeans (47)	37	10	0	37	10	0	.89

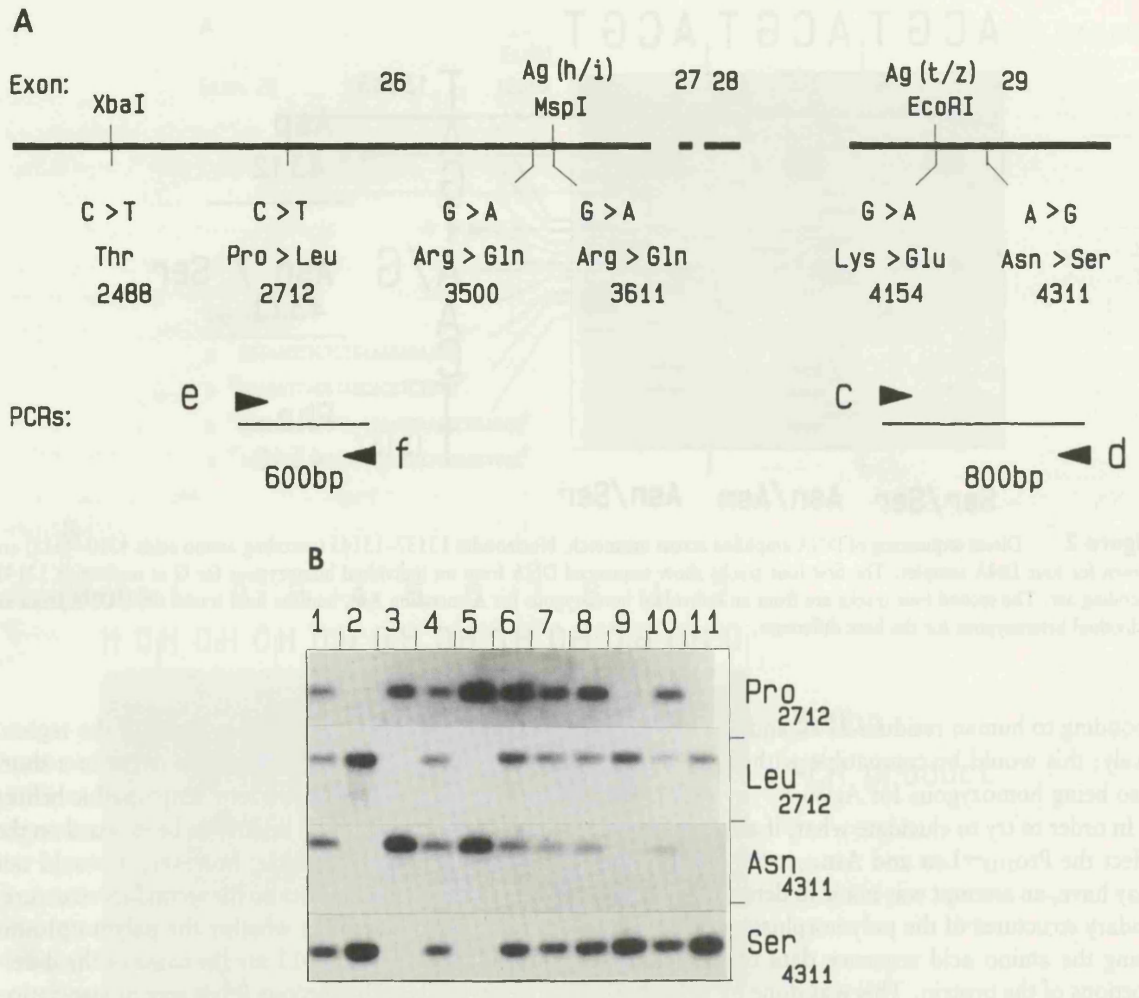


Figure 3 Allele-specific oligonucleotide melting to detect Pro_{2712}→Leu and Asn_{4311}→Ser. *Top,* Map of exons 26–29, showing positions of single base changes and amino acid substitutions that they create. Where the changes also generate RFLPs and antigen group (Ag) polymorphisms, these are shown along the top. The positions of the PCRs used to detect the base changes are shown below. The sequences of amplimers c and d are given in fig. 1a. The sequences, 5′ to 3′, of the other amplimers are as follows: e—ATCATCAGAACCATTGACCAGATGCTGAAAC; and f—TGACAATCACTCCATTACTAAGCTCCAGTG. The sequences, 5′ to 3′, of the ASOs are as follows: Pro—CACATACCAGAATTC; Leu—GAATTCTAGTATGTG; Asn—AATCTTCAATGATTATA; and Ser—TATAATCACTGAAGATT. *Bottom,* Autoradiograph showing example of results obtained. Quadruplicate blots of amplified DNA from 11 individuals are shown. These are hybridized to the ASOs recognizing the alleles for (from top to bottom) Pro_{2712}, Leu_{2712}, Asn_{4311}, Ser_{4311}. Sample 1 is an illustration of DNA from an individual heterozygous at both sites; sample 2 is from one homozygous for Leu_{2712}/Ser_{4311}; and sample 3 is from one homozygous for Pro_{2712}/Asn_{4311}.

leles are found unequivocally on both the X+ and X− alleles (table 3).

The effects that these different alleles have on the lipid, lipoprotein, and apolipoprotein parameters were compared in a sample of healthy Swedish individuals (tables 4 and 5). The *Xba*I RFLP accounts for as much as 9.1% of the phenotypic variance in serum LDL-cholesterol levels ($P < .05$; table 4) and for 8.9% of the variance in serum apo B levels, although this value does not reach statistical significance. In each case, the X− allele is associated with reduced mean levels of both LDL-cholesterol and apo B (table 5). By contrast, the Pro_{2712}→Leu/Asn_{4311}→Ser polymorphisms explain less of the variance seen in LDL-cholesterol and apo B levels—i.e., only 1.1% and 2.8%, respectively (table 4)—although in each case there is also a trend for the Leu_{2712}/Ser_{4311} allele to be associated with reduced mean levels (table 5). The polymorphisms at residues 2712 and 4311 do, however, explain a significant percentage of the variance

Table 2

Complete Association between Pro_{2712}→Leu, Asn_{4311}→Ser, and Ag(x/y)

Sample	Ag(x/y) xx	xy	yy	
Individuals from Finland:				
Pro_{2712}→Leu:				
LL	9			
LP		44		
PP			66	
Total				119
Asn^{4311}→Ser:				
SS	9			
SN		44		
NN			66	
Total				119
	Asn_{4311}→Ser			
	SS	SN	NN	
Combined ethnic groups:				
Pro_{2712}→Leu:				
LL	96			
PL		194		
PP			233	
Total				523

Table 3

Linkage Disequilibrium with Allelic Association between *Xba*I RFLP, Pro_{2712}→Leu, and Asn_{4311}→Ser in 420 Individuals from Different Populations

	*Xba*I X−X−	X−X+	X+X+
(SL)	65	0	0
(SP)	86	76	0
(NP)	26	93	74

Note.—Δ = .57; P < .001.

(9.0%, $P < .05$; table 4) seen in serum HDL-cholesterol levels. In this instance the Leu_{2712}/Ser_{4311} allele is associated with raised mean HDL-cholesterol levels (table 5).

Discussion

Our findings demonstrate complete allelic association between Ag(x/y) and polymorphism at sites encoding apo B residues 2712 and 4311 in the sample

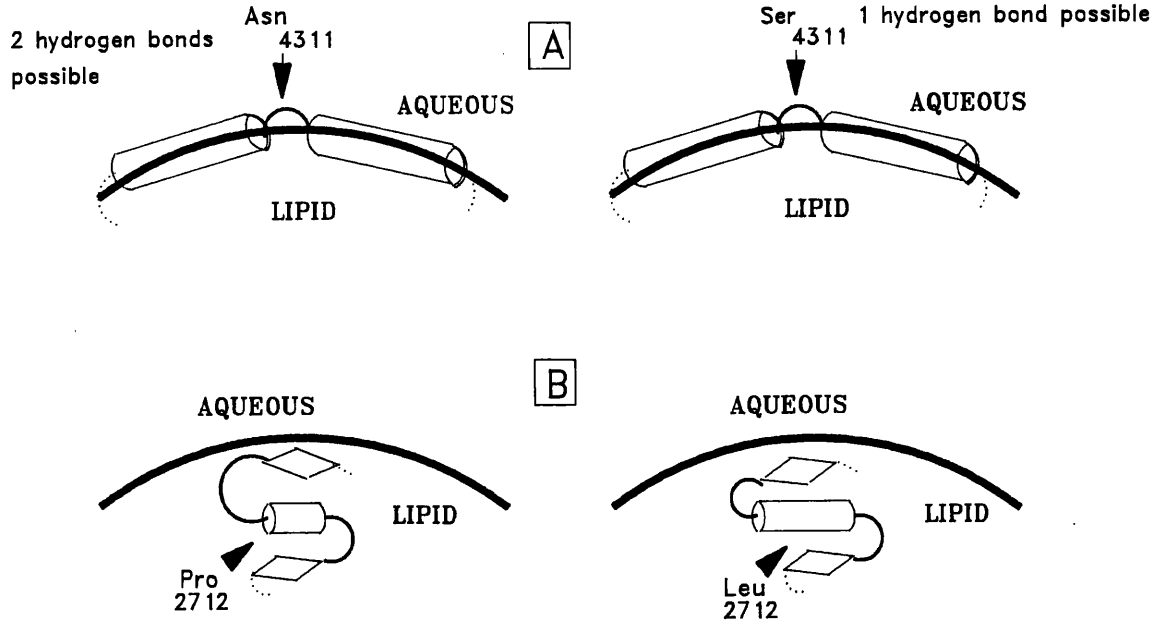


Figure 4 Cartoon representations of models of possible apo B structure in regions of amino acid substitutions. The phospholipid membrane of the LDL particle is represented by the thicker-lined arcs. Potential helical regions are denoted by cylinders, and likely β-sheet/β-turn regions are denoted by parallelograms. (Cartoons are not drawn to scale.) Ⓐ, Residue 4311. This residue probably lies in a short hydrophillic region between two longer amphipathic helices. The Asn→Ser change does not appear to alter the secondary structure; however, Asn could partake in the formation of two hydrogen bonds, and Ser could partake in only one during the formation of any tertiary structure. Ⓑ, Residue 2712. This residue is in a hydrophobic region generally consisting of potential β-sheet and β-turn; however, residue 2712 lies within a short helix which appears to be broken by the presence of Pro at this site but which is sustained by Leu.

Table 4

Percentage of Variance (R^2 × 100) Explained by *Xba*I and Pro_{2712}→Leu/Asn_{4311}→Ser Genotypes in 86 Swedish Individuals

	*Xba*I	Pro_{2712}→Leu and Asn_{4311}→Ser
Total cholesterol	5.4	1.5
Total triglyceride	1.5	1.1
LDL cholesterol	9.1*	1.1
HDL cholesterol	6.3	9.0*
HDL/LDL cholesterol	8.6*	7.8*
Apo B	8.9	2.8
Apo AI	2.2	1.8

* $P < .05$.

from Finland. Thus the epitopes recognized as Ag(x/y) could be generated by a single amino acid substitution at either residue 2712 or residue 4311. Alternatively, if both these residues are involved in the formation of the tertiary structure in the apo B protein, they may interact to form one epitope, with Ag(x) consisting of Leu_{2712} plus Ser_{4311} and Ag(y) of Pro_{2712} plus Asn_{4311}. Since there is little, if any, evidence of recombination in this region of the apo B gene, it is also possible that neither of these residues forms the epitope but that another, as yet undefined amino acid substitution, in complete linkage disequilibrium with these alleles, may create the Ag(x/y) epitopes. Recently, Wu et al. (1991) have published an association study, using 18 individuals protein phenotyped for the Ag polymorphisms, in which they conclude that the Pro_{2712}→Leu substitution is "responsible for the Ag(x/y) polymorphism." While our studies confirm that the alleles encoding residue 2712 are in complete allelic association with those encoding the Ag(x/y) epitopes, our findings demonstrate that such a conclusion may be unwarranted because the correlation seen between residue 2712 and Ag(x/y) is not proof of causation. The same caution must also be applied to the molecular bases of the other Ag polymorphisms, which have also been deduced from association studies. Investigations using in vitro expression of constructs with combinations of the amino acid variants together with antibodies which detect the different Ag epitopes are required to test which apo B residues actually form the Ag epitopes.

The secondary-structure protein modeling presented is inherently speculative but does give some indication as to which of the amino acid substitutions described here could be the molecular basis of the Ag(x/y) antigenic determinants. Pro_{2712} is within one of the hydrophobic "proline-rich" consensus sequences postulated to be able to penetrate the lipid of the LDL particle (DeLoof et al. 1987*b*) and is thus unlikely to be exposed on the surface of the LDL particle. DeLoof et al. (1987*b*) place Pro_{2712} as the ninth residue of a 25-residue consensus motif that occurs eight times within apo B100 (given in one-letter code, this consensus motif is FQVPDLHIPEFQLPHISHTIEVPTF). It is notable that in six of these eight motifs proline is conserved as the ninth residue and that in the other two it is replaced once by asparagine and once also by leucine, which indicates that the Pro→Leu change may be conservative within such a lipid-binding region. The Asn_{4311}→Ser change is also not predicted to have a major effect on apo B conformation, unless this residue is involved in hydrogen bonding, in which case Asn could participate in the formation of two bonds and Ser in only one. Residue 4311 occurs within a short joining region between two long amphipathic helices (fig. 4), and such helices are thought to have

Table 5

Comparison of Mean ± SE Serum Levels of Cholesterol and apo B by Genotype for *Xba*I RFLP and Pro_{2712}→Leu/Asn_{4311}→Ser Polymorphism in 83 Swedish Individuals

	*Xba*I			Pro_{2712}→Leu AND Asn_{4311}→Ser		
	X−X−	X−X+	X+X+	(LS)	(LN)	(PN)
No. of individuals	17	39	27	4	23	56
LDL cholesterol (mmol/liter)	3.57* ± .65	3.88 ± .83	4.25 ± .90	3.77 ± .89	3.85 ± .61	3.98 ± .93
HDL cholesterol (mmol/liter)	1.53 ± .38	1.33 ± .29	1.48 ± .29	1.85* ± .45	1.40 ± .35	1.40 ± .28
Apo B (mmol/liter)	99 ± 13	106 ± 18	111 ± 16	94 ± 14	106 ± 15	107 ± 18

* $P < .05$.

"surface-seeking" properties within the LDL particle (DeLoof et al. 1987*a*). This indicates that residue 4311 may be exposed on the surface of the particle and would therefore be available to act as an epitope for an antibody. Another possibility is that residues 2712 and 4311 are closely related in space on the surface of the lipoprotein particle, because of the folding of the protein. If this is so, then they could jointly form the Ag(x/y) epitopes, Asn_{4311} interacting with Pro_{2712} or Ser_{4311} with Leu_{2712} as the only two functional combinations of the protein. Although it is not yet possible to use computer modeling to explore the possibility that residues 2712 and 4311 may interact in the tertiary structure, such interaction cannot be excluded.

These data demonstrate strong linkage disequilibrium in this region of the apo B gene, with apparent complete allelic association between the Ag(x)/Leu_{2712}/Ser_{4311} and *Xba*I X− alleles in the 688 chromosomes where phase could be determined unambiguously. Berg et al. (1986) also found strong linkage disequilibrium between the Ag(x) and *Xba*I X− alleles in a sample of 75 Norwegian individuals, but they did not find complete allelic association; the reason for this difference remains unclear. Both sets of data indicate that the mutations creating the Ag(x)/Leu_{2712}/Ser_{4311} allele may have arisen on a chromosome carrying the X− allele, and our own data support the theory that they have not since been separated by recombination. Thus, our original aim of searching for common potentially functional variants of apo B, in linkage disequilibrium with the alleles of the *Xba*I RFLP, has been successful, although association studies alone cannot exclude or confirm that the Ag(x)/Leu_{2712}/Ser_{4311} allele is having a functional effect on apo B. The results of our initial analysis of the Swedish sample imply that the Ag(x)/Leu_{2712}/Ser_{4311} allele is associated with some, but not all, of the differences in both serum LDL-cholesterol and apo B levels seen in association with the *Xba*I X− allele. In addition, the same allele is associated with raised HDL-cholesterol levels. This pattern is consistent with the fact that the Leu_{2712}/Ser_{4311} alleles are a subset of the *Xba*I X− alleles. However, our data also predict that there are other, as yet undetected, functional variants of apo B that will account for the rest of the differences seen in association with the *Xba*I RFLP. Further detailed analyses in a much larger sample are required to address this question more fully.

Recently, Rapacz et al. (1991), using protein phenotyping, have reported that there is only one Ag haplotype—Ag(g,d,y,i,t)—in 20 chimpanzees and eight gorillas that they have studied. They thus suggest that this is the ancestral haplotype for human apo B. This also implies that the polymorphism, seen at each of the Ag loci in *Homo sapiens,* has evolved since the divergence of the ancestors of man and higher apes, estimated to be 5 million years ago. We genotyped DNA from four chimpanzees, four gorillas, and two orangutans and found them all to be homozygous for both proline and asparagine at the apo B residues corresponding to human residues 2712 and 4311, respectively. Thus these results support those of Rapacz et al., and it can be predicted that these apes may also have the Ag(y) epitope. However, since the complete association of Pro_{2712}/Asn_{4311} is also present in these apes, it is still not possible to determine whether either, both, or neither of these residues form the Ag(y) epitope.

Since Ag(x/y) and Ag(a1/d) are the only Ag loci that are polymorphic in every ethnic group worldwide, Breguet et al. (1990) suggest that these loci must be the oldest of the Ag variant sites and that they probably evolved before the postulated migration of *H. sapiens* out of Africa, estimated to have occurred approximately 100,000 years ago. If this is the case, then it indicates that there has been no recombination between the sites encoding Ag(x/y) and residues 2712 and 4311, a distance of approximately 7 kB of DNA, in the history of modern man. Analysis of a larger region of the apo B gene lends further credence to this. When our own data are used in conjunction with those of Breguet et al. (1990), it is possible to draw an evolutionary tree for the 3′ end of the apo B gene (fig. 5). The two haplotypes, designed A_0 and B_0 (fig. 5), differ by only the C-to-T base change that creates the *Xba*I polymorphic site. It is not possible to identify which of these chromosomes arose first, since both are present in all major ethnic groups; however, the frequencies of these two forms vary widely between populations. It is possible to deduce that all the other mutations arose on one or other of these chromosomes in a sequential order, and no subsequent recombination is required to explain the observed haplotypes. Haplotype A_1 can be postulated to have been created by a G-to-A base change on A_0, giving rise both to the *Eco*RI RFLP, at the DNA level, and to the Ag(t/z) polymorphism, at the protein level. The G-to-A mutation that generated Arg_{3500}→Gln (Innerarity 1990) (designated A_2 in fig. 5) appears to have arisen on an A_1 chromosome, since all the patients so far identified as having this disease have inherited this haplotype

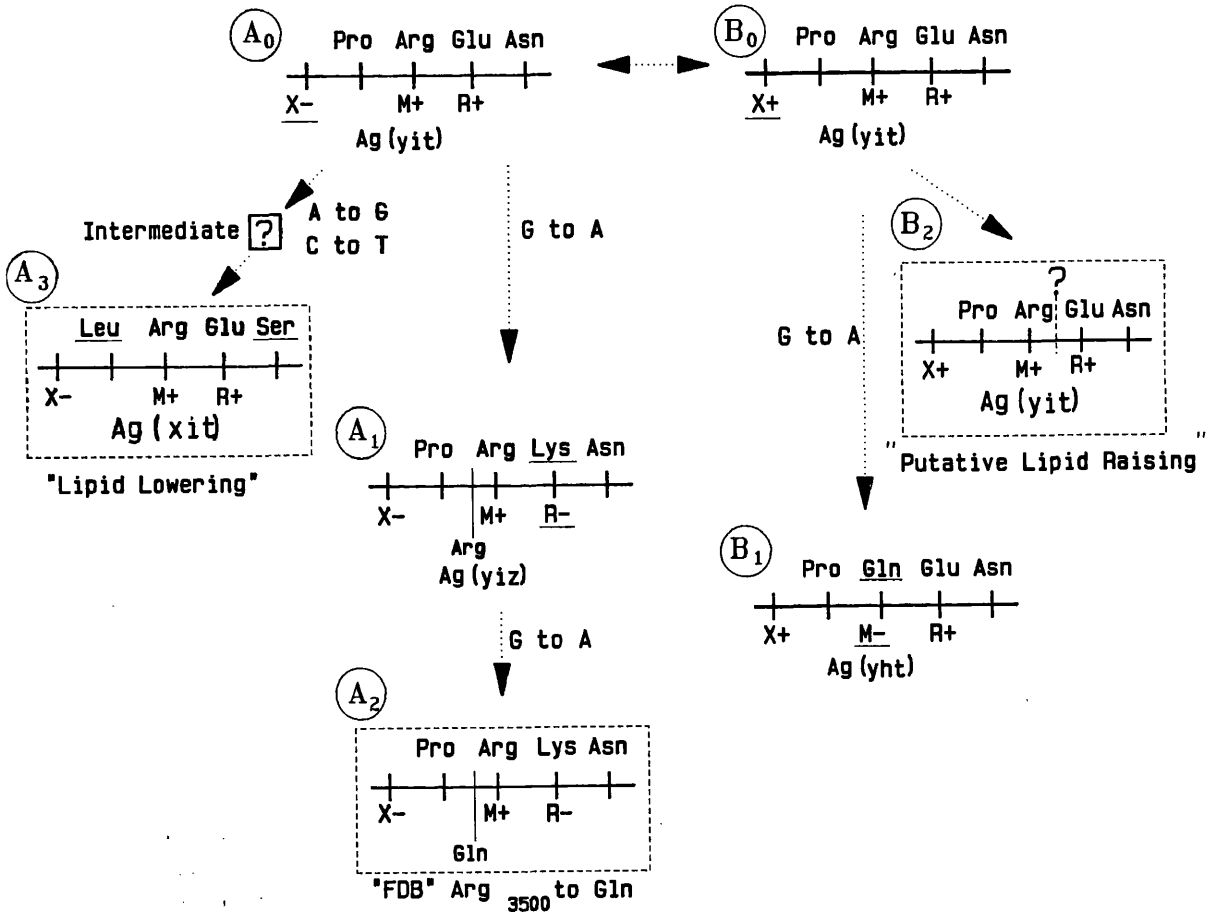


Figure 5 Postulated evolutionary tree for the 3′ end of apo B Gene. Line representations of the apo B haplotypes show amino acids above the line and show both the presence (+) or absence (−) of restriction-enzyme sites (X = *Xba*I; M = *Msp*I; and E = *Eco*RI) and partial Ag haplotypes beneath. Probable progenitor haplotypes (A_0 and B_0) are shown, and all sequential haplotypes have been given a number. Base changes creating the new haplotypes are shown next to the arrows. Full discussion is given in the text.

(Ludwig and McCarthy 1990; Myant et al. 1991). The only exception to the pattern that each haplotype differs from the previous one by a single base change is Ag(xit) (designated A_3 in fig. 5), which differs from A_0 at the two sites: Pro_{2712}→Leu and Asn_{4311}→Ser. Our studies have not revealed an intermediate haplotype (table 2), implying that these two base changes may have arisen within a very short time interval. Our data and those of Berg et al. (1976) suggest that this haplotype has a small lipid-lowering effect. We postulate that the Ag(h/i) polymorphism, the *Msp*I RFLP at the DNA level, has been created by a G-to-A base change on haplotype B_0 (results not shown), thereby giving rise to B_1. We also postulate that there is an additional functionally distinct haplotype (designated B_2) that has been created by an as yet unidentified sequence change occurring on the B_0 haplotype. This sequence change would be in allelic association with the *Xba*I X+ site and has the effect of raising serum cholesterol levels (Berg et al. 1986; Law et al. 1986; Talmud et al. 1987; Aalto-Setälä et al. 1989; Paulweber et al. 1990).

There is some evidence—e.g., the extremely rare haplotype Ag(xiz)—that recombination events may have occurred in this gene region. (Breguet et al. 1990; table 4). However, in general, the evidence indicates that novel mutation, rather than recombination, is the driving force for change in the 3′ end of the apo B gene. This unexpected result poses the question of whether this lack of recombination is due to chance or selection. In order for selection to be the cause, it would require that the forms of apo B created by recombina-

tion be dominant and lethal; these are not attributes that have been generally proposed for variants of the apo B gene. However, since introns have been postulated to be "spare" DNA available for mutation and recombination (Gilbert 1985), the presence of huge exons and only two small introns in the 3′ end of the apo B gene may indicate that recombination in the apo B is disadvantageous.

Acknowledgments

This work was supported by the Sunley Research Trust and by grants from the British Heart Foundation (grant RG5), the Sigoid Juselius Foundation, the Academy of Finland, the University of Helsinki, the Swedish Medical Research Council (grants 08691 and 08152), the Swedish Heart-Lung Foundation, the Torsten and Ragnar Söderberg Foundation, and the Singapore National Heart Association. H.-H.R. is supported by a grant from the Deutsche Forschungsgemeinschaft; R.P. is supported by the Medical Research Council; and C.-F.X. is supported by the British Council. The ape DNA samples were a gift from Dr. Ian Gray and Professor Alec Jeffreys. We would also like to thank Drs. Simon Butler and Alan Mills for helpful discussions and Elaine Osman for help with preparation of the manuscript.

References

Aalto-Setälä K, Gylling H, Helve E, Kovanen P, Miettinen TA, Turtola H, Kontula K (1989) Genetic polymorphism of the apolipoprotein B gene locus influences serum LDL cholesterol level in familial hypercholesterolaemia. Hum Genet 82:305–307

Allison AC, Blumberg BS (1961) An isoprecipitation reaction distinguishing human serum protein types. Lancet 1: 634–637

Berg K (1987) Genetics of coronary heart disease and its risk factors. In: Bock G, Collins GM (eds) Molecular approaches to human polygenic disease. Ciba Foundation Symposium, 130, 1st ed. Wiley & Sons, Chichester, pp 14–28

Berg K, Hames C, Dahlen G, Frick MH, Krishan I (1976) Genetic variation in serum low density lipoproteins and lipid levels in man. Proc Natl Acad Sci USA 73:937–940

Berg K, Powell LM, Wallis SC, Pease R, Knott TJ, Scott J (1986) Genetic linkage between the antigenic group (Ag) variation and the apolipoprotein B gene: assignment of the Ag locus. Proc Natl Acad Sci USA 83:7367–7370

Blackhart BD, Ludwig EH, Pieratti VR, Caiati L, Onasch MA, Wallis SC, Powell L, et al (1986) Structure of the human apolipoprotein B gene. J Biol Chem 261:15364–15367

Breguet G, Bütler R, Bütler-Brenner E, Sanchez-Mazas A (1990) A worldwide population study of the Ag-system haplotypes, a genetic polymorphism of human low-density lipoprotein. Am J Hum Genet 46:502–517

Carlsson P, Darnfors C, Olofsson SO, Bjursell G (1986) Analysis of the human apolipoprotein B gene: complete structure of the B-74 region. Gene 49:29–51

Chakravati A, Buetow KH, Antonarakis SE, Waber PG, Boehm CD, Kazazian HH (1984) Nonuniform recombination within the human β-globin gene cluster. Am J Hum Genet 36:1239–1258

Cotton RG, Rodrigues NR, Campbell RD (1988) Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proc Natl Acad Sci USA 85:4397–4401

De Loof H, Rosseneu M, Brasseur R, Ruysschaert JM (1987*a*) Functional differentiation of amphiphilic helices of the apolipoproteins by hydrophobic moment analysis. Biochim Biophys Acta 911:45–52

De Loof H, Rosseneu M, Yang C-Y, Li W-H, Gotto AM, Chan L (1987*b*) Human apolipoprotein B: analysis of internal repeats and homology with other apolipoproteins. J Lipid Res 28:1455–1465

Demant T, Houlston RS, Caslake MJ, Series JJ, Shepherd J, Packard CJ, Humphries SE (1988) Catabolic rate of low density lipoprotein is influenced by variation in the apolipoprotein B gene. J Clin Invest 82:797–802

Dunning AM, Houlston R, Frostegard J, Revill J, Nilsson J, Hamsten A, Talmud P, et al (1991) Genetic evidence that the putative receptor binding domain of apolipoprotein B (residues 3130 to 3630) is not the only region of the protein involved in interaction with the low density lipoprotein receptor. Biochim Biophys Acta 1096:231–237

Dunning AM, Tikkanen MJ, Enholm C, Butler R, Humphries SE (1988) Relationships between DNA and protein polymorphisms of apolipoprotein B. Hum Genet 78:325–329

Durrington PN, Hunt L, Ishola M, Kane J, Stephens WP (1986) Serum apolipoproteins AI and B and lipoproteins in middle aged men with and without previous myocardial infarction. Br Heart J 56:206–212

Ehnholm C, Huttunen JK, Retinen P, Leine LI, Mutanen M, Kostiainen E, Iacono JM, et al (1984) Effect of a diet low in saturated fatty acids on plasma lipid, lipoprotein and HDL subfraction. Arteriosclerosis 4:265–269

Ehnholm C, Huttunen JK, Rietinen P, Leino M, Mutanen M, Kostiainen E, Pikkarainem J, et al (1982) Effect of a diet on serum lipoproteins in a population with a high risk of coronary disease. N Engl J Med 307:850–855

Eisenberg D (1984). Three dimensional structure of membrane and surface proteins. Annu Rev Biochem 53:595–623

Frostegård J, Hamsten A, Gudland M, Nilsson J (1990) Low density lipoprotein–induced growth of U937 cells—a

novel method to determine the receptor binding of low density lipoprotein. J Lipid Res 31:33–34

Gilbert W (1985) Genes-in-pieces revisited. Science 228: 823–828

Hamsten A, Eféndic S, Walldius G, Szamosi A, de Faire U (1987) Glucose tolerance and insulin response to glucose in nondiabetic young male survivors of myocardial infarction. Am Heart J 113:917–927

Hamsten A, Iselius L, Dahlén G, de Faire U (1986*a*) Genetic and cultural inheritance of serum lipids, low and high density lipoprotein cholesterol and serum apolipoproteins A-I, A-II and B. Atherosclerosis 60:199–208

Hamsten A, Walldius G, Dahlén G, Johansson B, de Faire U (1986*b*) Serum lipoproteins and apolipoproteins in young male survivors of myocardial infarction. Atherosclerosis 59:223–235

Hasstedt SJ, Wu L, Williams RR (1987) Major locus inheritance of apolipoprotein B in Utah pedigrees. Genet Epidemiol 4:67–76

Houlston RS, Turner PR, Revill J, Lewis B, Humphries SE (1988) The fractional catabolic rate of low density lipoprotein in normal individuals is influenced by variation in the apolipoprotein B gene. Atherosclerosis 71:81–85

Huang LS, Gavish D, Breslow JL (1990) Sequence polymorphism in the human apo B gene at position 8344. Nucleic Acids Res 18:5922

Innerarity TL (1990) Familial hypobetalipoproteinemia and familial defective apolipoprotein B100: genetic disorders associated with apolipoprotein B. Curr Opinions Lipidol 1:104–109

King-Underwood L, Gudnason V, Humphries S, Seed M, Patel D, Knight B, Soutar A (1991) Identification of the 664 proline to leucine mutation in the low density lipoprotein receptor in four unrelated patients with familial hypercholesterolaemia in the UK. Clin Genet 40:17–28

Kuusi T, Ehnholm C, Huttunen JK, Kostiainen E, Pietinen P, Leino U, Uusitalo U, et al (1985) Concentration and composition of serum lipoproteins during a low-fat diet at two levels of polyunsaturated fat. J Lipid Res 26:360–367

Law A, Wallis SC, Powell LM, Pease RJ, Brunt H, Priestley LM, Knott TJ, et al (1986) Common DNA polymorphism within coding sequence of apolipoprotein B gene associated with altered lipid levels. Lancet 1:1301–1303

Ludwig EH, McCarthy BJ (1990) Haplotype analysis of the human apolipoprotein B mutation associated with familial defective apolipoprotein B100. Am J Hum Genet 47: 712–720

Ma Y, Schumaker V, Butler R, Sparkes RS (1987) Two DNA restriction fragment length polymorphisms (RFLPs) associated with Ag (t/z) and Ag (c/g) antigenic sites of human apolipoprotein B. Arteriosclerosis 3:301–305

Ma Y, Wang X, Butler R, Schumaker VN (1989) Bsp 12861 restriction fragment length polymorphism detects AG (c/g) locus of human apolipoprotein B in all 17 persons studied. Arteriosclerosis 9:242–246

McKeigue PM, Shah B, Marmot MG (1991) Central obesity and insulin resistance underlie high diabetes prevalence and cardiovascular risk in South Asians. Lancet 337:382–386

Maniatis T, Fritsch EF, Sambrook J (eds) (1982) Molecular cloning: a laboratory manual, 1st ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Miller GJ (1989) Ethnicity, lipoproteins and haemostatic factors. In: Cruickshank JK, Beevers DG (eds) Ethnic factors in health and disease, 1st ed. Wright, London, pp 280–288

Montandon AJ, Green PM, Giannelli F, Bentley DR (1989) Direct detection of point mutations by mismatch analysis: application to haemophilia B. Nucleic Acids Res 17: 3347–3358

Myant NB, Gallagher JJ, Knight BL, McCarthy SN, Frostegård J, Nilsson J, Hamsten A, et al (1991) Clinical signs of familial hypercholesterolaemia in patients with familial defective apolipoprotein B-100 and normal low-density lipoprotein receptor function. Atherosclerosis Thromb 11:691–703

Pairitz G, Davignon J, Mailloux H, Sing CF (1988) Sources of interindividual variation in the quantitative levels of apolipoprotein B in pedigrees ascertained through a lipid clinic. Am J Hum Genet 43:311–321

Paulweber B, Friedl W, Krempler F, Humphries SE, Sandhofer F (1990) Association of DNA polymorphism at the apolipoprotein B gene locus with coronary artery disease and serum very low density lipoprotein levels. Arteriosclerosis 10:17–24

Rapacz J, Chen L, Butler-Brunner E, Wu M-J, Hasler-Rapacz J, Butler R, Schumaker V (1991) Identification of the ancestral haplotype for apolipoprotein B suggests an African origin of *Homo sapiens sapiens* and traces their subsequent migration to Europe and the Pacific. Proc Natl Acad Sci USA 88:1403–1406

Rosseneu M, De Loof H, De Meutter J, Ruysschaert JM, Vanloo B, Brasseur R (1990) Identification and computer modelling of functional domains in plasma apolipoproteins. In: Brasseur R (ed) Molecular description of biological membrane components by computer aided conformational analysis, vol. 2, 1st ed. CRC, New York, pp 173–190

Saha N, Tay JSH, Humphries SE. Apolipoprotein B gene DNA polymorphisms (XbaI and EcoRI) serum lipids and apolipoprotein B in healthy Chinese (submitted)

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, et al (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Talmud PJ, Barni N, Kessling AM, Carlsson P, Darnfors C, Bjursell G, Galton D, et al (1987) Apolipoprotein B gene variants are involved in the determination of serum cholesterol levels: a study in normo- and hyperlipidaemic individuals. Atherosclerosis 67:81–89

Tikkanen MJ, Xu C-F, Hamalainen T, Talmud P, Sarna S,

Huttunen JK, Pietinen P, et al (1989) XbaI polymorphism of the apolipoprotein B gene influences plasma lipid response to diet intervention. Clin Genet 37:327–334

Wang X, Schlapser P, Ma Y, Butler R, Elovson J, Schumaker VN (1988) The ag(a1/d) immunogenetic polymorphism coincides with a T to C substitution at nucleotide 1981, creating an AluI restriction site. Arteriosclerosis 8:429–435

Whayne TF, Alaupovic P, Curry MD, Lee ET, Anderson PS, Schechter E (1981) Plasma apolipoprotein B and VLDL-, LDL- and HDL-cholesterol as risk factors in the development of coronary artery disease in male patients examined by angiography. Atherosclerosis 39:411–424

Wu M-J, Butler E, Butler R, Schumaker VN (1991) Identification of the base substitution responsible for the Ag(x/y) polymorphism of apolipoprotein B-100. Arteriosclerosis Thromb 11:379–384

Xu CF, Nanjee N, Tikkanen MJ, Huttunen JK, Pietinen P, Butler R, Angelico F, et al (1989) Apolipoprotein B amino acid 3611 substitution from arginine to glutamine creates the Ag (h/i) epitope: the polymorphism is not associated with differences in serum cholesterol and apolipoprotein B levels. Hum Genet 82:322–326

Yang CY, Chen SH, Gianturco SH, Bradley WA, Sparrow JT, Tanimura M, Li WH, et al (1986) Sequence, structure, receptor-binding domains and internal repeats of human apolipoprotein B-100. Nature 323:738–742

Yang C-Y, Gu Z-W, Weng S-A, Kim TW, Chen S-H, Pownall HJ, Sharp PM, et al (1989) Structure of apolipoprotein B-100 of human low density lipoproteins. Arteriosclerosis 9:96–108

Young SG, Hubl ST (1989) An ApaLI restriction site polymorphism is associated with the MB19 polymorphism in apolipoprotein B. J Lipid Res 30:443–449

Biochimica et Biophysica Acta, 1165 (1993) 271–278

BBALIP 54062

A postulated phylogenetic tree for the human apolipoprotein B gene: unpredicted haplotypes are associated with elevated apo B levels

Alison M. Dunning [a], Helmut-H. Renges [a,1], Anders Hamsten [b], Philippa Talmud [a] and Steve Humphries [a]

[a] *Cardiovascular Genetics, Department of Medicine, University College and Middlesex School of Medicine, Rayne Institute, University Street, London (UK) and* [b] *King Gustaf V Research Institute, Karolinska Institute and Department of Internal Medicine, Karolinska Hospital, Stockholm (Sweden)*

(Received 16 March 1992)
(Revised manuscript received 20 August 1992)

Key words: *Xba*I RFLP; Antigen Group polymorphism; Haplotype; Low-density lipoprotein

Using published data on seven polymorphic sites in the human apolipoprotein B (apo B) gene, it is possible to postulate a model phylogenetic tree for this gene, covering the time since the divergence of human beings from other primates. This simple model assumes no obligatory recombination events or multiple occurrences of the same mutation. This model was tested in two samples of Swedish individuals consisting of 143 young, myocardial infarction patients and 90 healthy, age-matched, control individuals. All the haplotypes postulated in the simple model were observed unequivocally. However, in addition, three unpredicted haplotypes were unambiguously observed and a further nine, much rarer haplotypes were deduced to occur in these samples. The frequencies of the haplotypes postulated in the model do not differ between the patient and control samples, however most of the unpredicted haplotypes occur more frequently in the patient group than in the controls. Two of these unpredicted haplotypes, defined by the combination of the Antigen group (a) epitope and the presence of the *Xba*I cutting site, were associated with raised serum apo B levels in the control group and significantly elevated levels in the patient group. We propose that these observations explain in part the consistent association reported between the *Xba*I polymorphic site in the apo B gene and levels of plasma lipids.

Introduction

Mature apolipoprotein B100 (apo B) is a 4536 amino acid protein secreted by the liver as a constituent of very-low-density lipoprotein (VLDL). During the metabolism of VLDL, other apoproteins and triacylglycerols are removed, leaving apo B as the sole protein component of low-density lipoprotein (LDL). As such it has a central role in lipid metabolism, serving both to maintain the integrity of the particle [1,2] and mediating the clearance of LDL-cholesterol from the plasma via the LDL-receptor for which apo B is a ligand. Since elevated plasma levels of LDL-cholesterol are one of the recognised risk factors for the development of coronary heart disease [3], the identification of the factors that determine plasma LDL-cholesterol levels are of major public health importance. Twin studies and path analysis have demonstrated that genetic variation has a significant impact on both plasma levels of apo B and LDL-cholesterol with a heritability of 0.5–0.6 [4,5]. Family studies using complex segregation analysis have found evidence for the plasma levels of apo B being determined by a major gene with a contribution from environmental factors and genes of small or intermediate effect [6,7].

Extensive searches have been made for variant forms of apo B that could affect its function in lipid and lipoprotein metabolism. Variant forms of apo B can be broadly classified into two groups: common polymorphisms with a small effect on serum lipid levels in the individual but a significant impact on the mean levels in the population due to their high frequency; and rare mutations, such as Familial Defective apo B100 [8] which may have major effects on serum lipid levels in affected individuals, but a lesser role in terms of the whole population. Protein polymorphisms of apo B

Correspondence to: A.M. Dunning, Cardiovascular Genetics, Department of Medicine, University College and Middlesex School of Medicine, Rayne Institute, University Street, London, WC1E 6JJ, UK.

[1] Present address: H.-H. Renges, I. Medizin. Klinik, Universitäts-Krankenhaus Epperdorf, Martinstr. 52, D-2000 Hamburg 20, Germany.

were first noted as the Antigen Group [Ag] system of variants, detected using antisera from multiply transfused patients [9]. There are five reported pairs of epitopes, each pair representing alleles of the apo B gene, termed: Ag(a_1/d), Ag(c/g), Ag(h/i), Ag(t/z) and Ag(x/y) (reviewed in Ref. 10). Sequencing studies have revealed base changes generating amino acid substitutions that are candidates for being the molecular bases for the Ag epitopes (Fig. 1, 11–18) although none of these are yet confirmed by expression studies.

Strong linkage disequilibrium has been detected between many of these sites [13,16,19,20] demonstrating an evolutionary relationship between them, but the exact progression has not previously been deduced. Rapacz and co-workers [21] found that a group of chimpanzees and gorillas have only one Antigen Group haplotype (Ag(g,d,y,i,z,), which corresponds to haplotype 1, Fig. 1) and from this they deduced that this is the ancestral haplotype from which all the others have evolved. Using this as a starting point and a combination of our own data and that from Breguet and co-workers [10], we have previously postulated a phylogenetic tree based on polymorphism data for the 3′ end of the apo B gene (exons 26–29) which was consistent with a simple model that all mutations arose in a sequential order, and that required no obligatory recombination events to explain the observed haplotypes [18,22]. In this paper, again using a combination of our own data and that of Breuget et al. [10], we have extended this tree to include the polymorphisms found in the 5′ end of the gene (signal peptide Insertion (I)/Deletion (Δ), Thr-71 to Ile and Ala-591 to Val).

There have now been numerous reports of association between the *Xba*I RFLP and differences in serum LDL cholesterol and apo B levels, the presence of the cutting site (the X + allele) being associated with raised levels [23–27]. However the base change that creates the variant *Xba*I site cannot be the direct cause of the effects seen, since it is in the third base (wobble position) of codon 2488 and does not alter the threo-

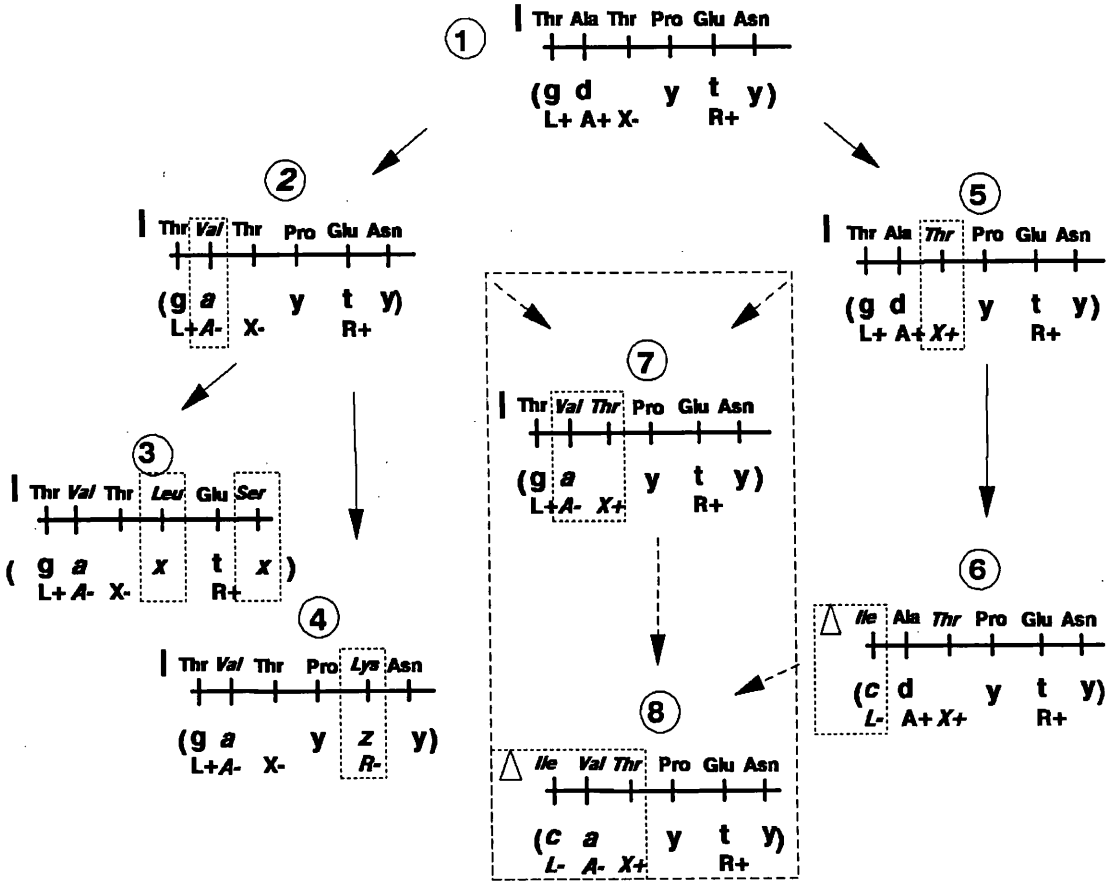


Fig. 1. Postulated phylogenetic tree for the apo B gene. The most parsimonious tree is shown around the outer edge (haplotypes 1–6). Haplotype 1 is the deduced ancestral haplotype [21]. Within the box are shown the two most frequently unambiguously observed complex haplotypes (haplotypes 7 and 8), together with arrows showing from which other haplotypes they may be derived. Amino acids (three letter codes) at residues 71, 591, 2488, 2712, 4154 and 4311, respectively, are shown above the line. The signal peptide Insertion/Deletion polymorphism is represented by I and Δ. The epitopes of the Ag polymorphisms for which the amino acids are candidate molecular bases are shown below the line, and the presence (+) or absence (−) of RFLP sites below these. (L = ApaLI, A = *Alu*I, X = *Xba*I, R = *Eco*RI). The differences creating each new haplotype are shown underlined.

nine residue at this position [28]. The changes in lipid parameters seen must be due to another relatively frequent base change or changes, that are in strong linkage disequilibrium with the alleles of this *Xba*I RFLP and have the potential to affect the function or expression of the protein. The direct causes of the altered LDL cholesterol levels are likely to be due either to altered synthesis rates of apo B and hence VLDL, or else changes in the catabolic rate of LDL mediated via the binding of apo B to the LDL receptor. In our originally proposed phylogenetic tree (Ref. 18, Fig. 6) the *Xba*I X + allele was observed on three distinct haplotypes. This raised the question of whether one of these three known haplotypes was associated with raised apo B and LDL cholesterol levels, or whether there was another, as yet undiscovered, variant form that had evolved on this X + branch of the tree.

In this study, we have determined the genotypes of and assigned haplotypes for 90 unrelated, normal Swedish individuals and 143 unrelated young Swedish myocardial infarction patients, using each of the polymorphic sites shown in Fig. 1, to investigate whether they are compatible with such an evolutionary tree and have examined the effects of the different haplotypes on serum lipid and lipoprotein parameters.

Materials and Methods

Patient and control groups

The patient group consisted of 143 male and female individuals from Stockholm county, Sweden, who had suffered a myocardial infarction below the age of 45 years. The control group consisted of 90 healthy, age-matched male residents of the same district, free of symptoms and clinical signs of coronary heart disease. Recruitment details, data on lipid, lipoprotein, apoprotein, and metabolic variables and descriptions of DNA preparation for these individuals have been presented elsewhere [29–31]. Individuals with serum cholesterol levels > 9.5 mM, and/or triacylglycerol levels > 3 mM, or with a diagnosis of familial hypercholesterolaemia, diabetes or porphyria were excluded from the lipid and lipoprotein analyses.

Detection of DNA polymorphisms

Genotyping of these individuals for the signal peptide Insertion/Deletion (I/Δ) polymorphism, Pro_{2712} to Leu, Asn_{4311} to Ser (Ag(x/y)), Glu_{4154} to Lys (Ag(t/z)), and the *Xba*I RFLP have been described previously [18,22]. However, reanalysis of these samples has lead to genotyping changes to two of the control individuals and 5 of the patients, causing slight differences in the genotype frequencies between the previous papers and this one.

Thr-71 to Ile (Ag(g/c)) was detected as an *Apa*LI RFLP [12]. Exon 4 of the apo B gene was amplified using oligonucleotides situated in introns 3 and 4 [32] according to the method of Saiki et al. [33] (Perkin Elmer/Cetus, CT USA) to give a 373 bp fragment. Under manufacturer's conditions, the enzyme *Apa*LI (Bioexcellance, Cambridge, UK) cuts the sequence GTGCAC, which encodes Thr-71, to give two fragments of 200 bp and 173 bp. The C to T base change which generates the codon for Ile_{71} also destroys the *Apa*LI recognition site. The DNA fragments were visualised on 2% Agarose gels (Gibco BRL, Paisley, UK).

Val_{591} to Ala (Ag(a/d)) was detected as an *Alu*I RFLP [14]. Oligonucleotide primers were used to amplify 82 bp of intron 13 together with the 5′ end of exon 14 [32] (Perkin Elmer/Cetus, CT USA) generating a 223 bp fragment. This fragment was digested with the enzyme *Alu*I (Bioexcellance, Cambridge, UK) under manufacturers instructions to give two fragments of 60 bp and 167 bp in alleles that encode Val_{591}, or three fragments of 60 bp, 50 bp and 117 bp in alleles that encode Ala-591. The DNA fragments were visualised on 2% Agarose gels (Gibco BRL, Paisley, UK).

Statistical analyses

Allele frequencies were estimated by the 'gene-counting' method and then the genotypes were tested for Hardy-Weinberg equilibrium by chi-square tests. Linkage disequilibrium between pairs of allelic sites was estimated using the correlation coefficient Δ [34]. The percentage of variance ($R^2 \times 100$) explained by different factors was estimated by linear regression using the SPSS/PC + computer program. Traits that were not already normally distributed were log transformed prior to this analysis. In the control group, age and body-mass-index (BMI) together accounted for 4.5% of the total variance in apo B levels ($P = 0.14$). In the patient group sex, age and BMI together accounted for 6.1% of the variance in apo B levels ($P = 0.09$). After adjustment for sex, age and BMI the impact of the apo B genotypes on apo B levels were estimated by non-linear regression and the differences between the means examined by one-way analysis of variance. If the differences between mean apo B levels were found to be significant, differences between LDL cholesterol and total cholesterol levels were also investigated by the same method. Apo B levels were plotted as percentiles for the control and patient groups separately, and then individuals carrying haplotypes 7 and 8 were identified on the plots. Chi-squared tests were used to investigate the distribution of these haplotypes.

Results

In this study seven polymorphic sites were examined: a 3 amino acid deletion in the signal peptide (Δ), Thr-71 to Ile (Ag(c/g)), Ala-591 to Val (Ag(a/d)),

TABLE I

Rare ellele frequencies for the polymorphisms studied in the patient and control groups

	Signal peptide ins/deletion Δ	Ag(c/g) Ag(c) = Ile_{71}	Ag(a/d) Ag(a) = Val_{591}	XbaI X-	Ag(x/y) Ag(x) = Leu_{2712}/Ser_{4311}	Ag(t/z) Ag(z) = Lys_{4154}
Controls	0.31	0.30	0.55	0.45	0.18	0.18
No. of Individuals	(90)	(87)	(90)	(90)	(90)	(90)
Patients	0.38	0.37	0.45 *	0.47	0.19	0.16
No. of Individuals	(144)	(142)	(142)	(143)	(141)	(143)

* $P < 0.05$. 95% CI for difference = 0.007–0.193.

Pro_{2712} to Leu, Asn_{4311} to Ser (Ag(x/y)), Glu_{4154} to Lys (Ag(t/z)), and the *Xba*I Restriction Fragment Length Polymorphism (RFLP) which does not alter residue Thr-2488. The proposed phylogenetic tree is shown in Fig. 1. In most cases, each haplotype appears to have been generated by a single base change occurring on the preceding haplotype, giving rise to a single amino acid substitution. However, there are two instances where there appear to be exceptions to this mechanism. The unique haplotype IgaX-xt (haplotype 3, Fig. 1) differs from its proposed progenitor IgaX-yt (haplotype 2, Fig. 1), at two sites - one substitution of Leu for Pro at residue 2712 and a second of Ser for Asn at residue 4311–with no haplotype representative of any intermediate stage being observed [18]. This suggests that either these two base changes occurred simultaneously or, that there was an intermediate haplotype which due to chance or selection is now extremely rare in the human gene pool sampled to date. The other instance is the unique haplotype ΔcdX + yt (haplotype 6, Fig. 1) which differs from its proposed precursor, IgdX + yt (haplotype 5, Fig. 1) by the substitution of Ile for Thr at residue 71 and also by the three amino acid deletion in the signal peptide (Δ). Although there are some individuals with haplotypes carrying a Δ allele in conjunction with an allele encoding Ag(g), or I with Ag(c) (Table II, present study, and Ref. 35), which may indicate an intermediate haplotype, the frequency of such chromosomes in the samples examined to date is low.

The two groups of individuals, patients and controls, were genotyped for all the polymorphisms considered in this study and the frequencies of each of the polymorphisms are shown in Table I. The genotype frequencies of all the polymorphisms were as expected for a sample in Hardy-Weinberg equilibrium. The allele

TABLE II

Unequivocally observed and assigned haplotypes in the patient and control groups

Differences in haplotype frequencies between control and patient groups: Haplotypes 7 and 8, 27/180 vs. 27/286, Chi-square = 3.25, $P < 0.1$. Haplotypes 9 to 18, 3/180 vs. 21/286, Chi-square = 7.30, $P < 0.01$.

	Controls		Patients		Total
	unequivocal	assigned	unequivocal	assigned	
1 Igd X−yt	1	8	11	11	30
2 Iga X−yt	2	3	1	6	13
3 Iga X−xt	13	19	7	38	76
4 Iga X−yz	10	23	7	35	75
5 Igd X+yt	10	12	21	11	54
6 Δcd X+yt	19	30	48	42	135
7 Iga X+yt	8	15	9	16	46
8 Δca X+yt	3	1	0	2	6
9 Δcd X+yt	0	1	4	5	10
10 Δca X+yz	0	0	0	3	3
11 Δcd X−yz	0	0	0	1	1
12 Δcd X−xt	0	0	0	1	1
13 Δcd X+xt	0	0	0	2	2
14 Δga X+yt	0	1	0	0	1
15 Δgd X+yt	0	0	0	1	1
16 Δga X−xt	0	1	0	2	3
17 Icd X+yt	0	0	0	1	1
18 Igd X−	0	0	0	0	11
Total:	66	114	108	178	466

frequencies were not significantly different between patients and controls for all of the polymorphisms with the exception of Ala-591 to Val (Ag(a/d)), where the frequency of Val is higher in the controls than in the patients. The degrees of pairwise linkage disequilibria between each of the polymorphisms are shown in Fig. 2. There is significant disequilibrium ($P < 0.05$) between all the sites considered.

The genotypes of all the individuals were assigned to haplotypes, commencing with those which could be determined unambiguously in the individuals who were homozygous at either all of the polymorphic sites or all except one. The haplotypes that were assigned in this manner fell into nine groups (Table II, haplotypes 1–9). Six of these nine types (haplotypes 1–6) corresponded to the haplotypes that we had predicted to exist in the parsimonious phylogenetic tree model–assuming no recombination events between sites. 82% of all the haplotypes characterised were classified into these haplotypes 1–6. However, in order to explain the other three unambiguously determined haplotypes: gaX + yt, ΔcaX + yt and ΔcdX-yt (haplotypes 7–9), it is necessary to propose either the occurrence of recombination events within the apo B gene, or some other mechanism.

The haplotypes of the remaining individuals were assigned by fitting their genotypes firstly to the six haplotypes proposed in the most simple model and secondly to the other three unequivocally characterised haplotypes. This left 2/90 (2.2%) individuals from the control group and 14/144 (9.7%) from the patient group whose haplotypes could not be assigned to groups that had already been demonstrated to occur unambiguously. For these individuals, an attempt was made to assign one of the individual's two chromosomes to a known haplotype leaving another previously unobserved one, these are shown in Table II (haplotypes 10–18). Although this system necessarily involves estimation and therefore inherent errors, three of the eight haplotypes deduced in this manner apparently occur more than once (haplotypes 10, 13 and 16) suggesting that they may be real. It is notable that the frequencies of these assigned haplotypes (haplotypes 9–18) are significantly higher in the patient group (21/286 = 0.073) than in the control group (3/180 = 0.017) (Table II).

The effects of the *Xba*I RFLP on serum apo B levels were investigated and the results are shown in Table III. We have previously reported [31] that in these samples the alleles of the *Xba*I RFLP are associated with a significant effect on apo B levels in the group of control individuals, with a similar trend on LDL cholesterol levels, the presence of the *Xba*I cutting site being associated with higher mean levels of

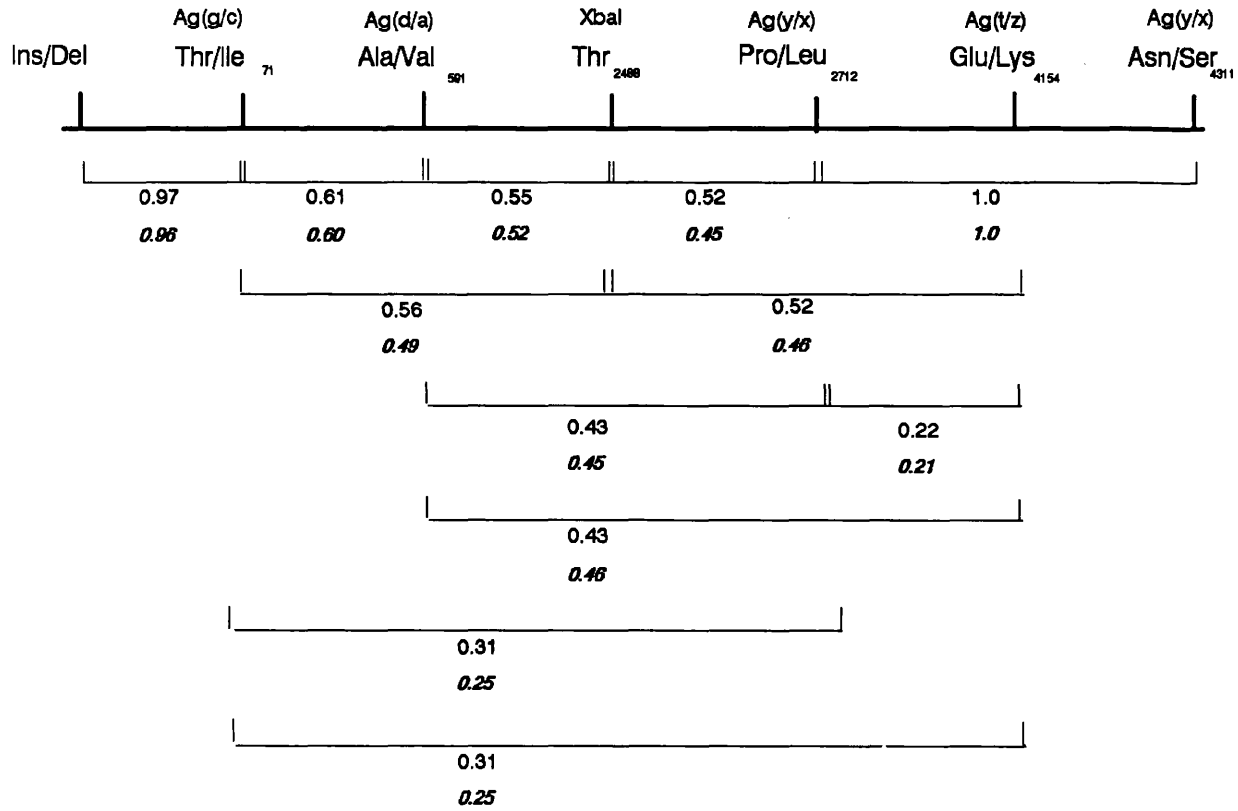


Fig. 2. Pairwise linkage disequilibria for each of the polymorphisms studied. Correlation coefficients (Δ) are shown for all the pairs. The upper figure is that for the control group, and the lower figure (in italics) for the patient group. All values are significantly higher than zero ($P < 0.05$).

both traits. None of the alleles of the other polymorphisms, considered individually, were associated with significant differences in apo B levels in either the patient or control groups. Two of these alleles, the Ag(x) and Ag(z), uniquely define single haplotypes (haplotypes 3 and 4, Fig. 1) and so these two haplotypes may be discounted from being associated with differences in apo B levels in these Swedish population samples. Similarly the signal peptide deletion allele (Δ) and the Ag(c) allele, are predominantly associated with haplotype 6 and so it can also be deduced that haplotype 6 is not associated with significantly different mean apo B levels significantly different from the other haplotypes (results not shown).

Since the existence of three of the unequivocally observed haplotypes had not been predicted in the most simple model phylogenetic tree, we next investigated the effects of these haplotypes on plasma apo B levels. The haplotype defined by the presence of Δ and Ag(c) in conjunction with the *Xba*I X − allele (haplotype 9, ΔcaX-yt) occurred too rarely (1/180 in the control group and 9/286 in the patient group) for statistical analysis. However the other two haplotypes, which are both defined by the presence of Ag(a) in combination with the *Xba*I X + allele (haplotype 7, IgaX + yt and haplotype 8, ΔcaX + yt) have combined frequencies of 27/180 (0.15) in the control and 25/286 (0.09) in the patient group. When the mean apo B levels in the groups of individuals with and without these haplotypes were compared, haplotypes 7 and 8 were associated with significantly higher mean apo B levels in the patient group and with a similar trend in the control group (Table III). Individuals with these haplotypes had also raised LDL cholesterol levels in both the patient and control groups but these differences did not reach statistical significance (results not shown). Individuals with haplotypes 7 and 8 were compared by quartiles of apo B levels (Fig. 3) and there are

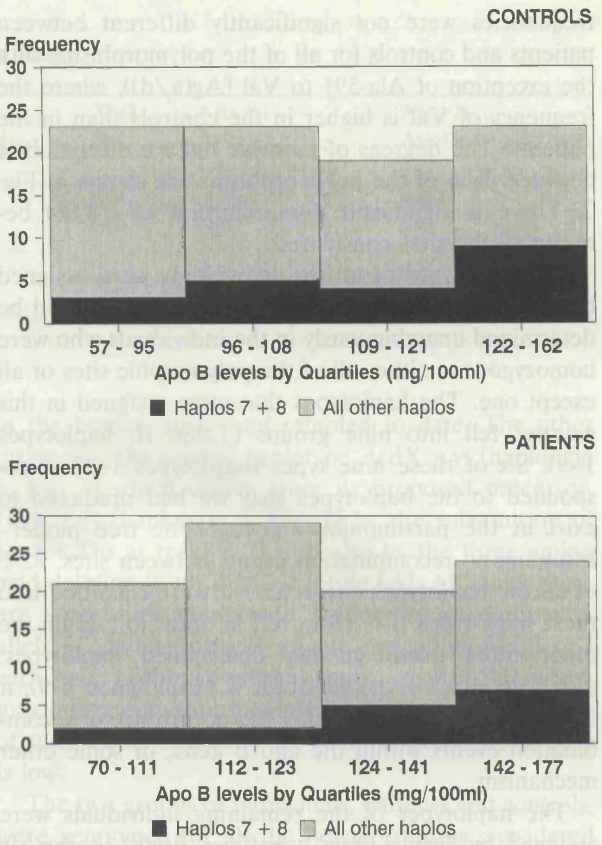


Fig. 3. Analysis of apo B levels within the two groups by percentiles. Black boxes represent individuals carrying one or two copies of haplotypes 7 and 8 and grey boxes represent individuals with all other haplotypes.

a higher than expected frequency of individuals with these haplotypes in the upper quartiles, 42% (9/21) of individuals with haplotypes 7 and 8 in the control group have apo B levels above the 75th percentile ($P < 0.05$), and in the patient group 44% (7/16) have apo B levels above the 75th percentile ($P < 0.1$). This confirms the view that, in both patient and control groups, these haplotypes are associated with elevated plasma apo B levels.

TABLE III

Mean apo B levels mg / dl (± S.E.) in patients and controls grouped by XbaI genotype or presence / absence of haplotypes 7 and 8

*Xba*I genotype	Controls		Patients	
	No.	Mean ± S.E.	No.	Mean ± S.E.
X − X −	18	99.5 ± 3.2	24	127.4 ± 3.2
X − X +	40	109.7 ± 3.3	43	120.6 ± 3.5
X + X +	25	112.3 ± 3.2	26	128.3 ± 4.6
	$F = 2.84$		$F = 1.35$	
	$P = 0.06$		$P = 0.27$	
Haplotypes 7 and 8				
Haplotypes 7 and 8	21	113.3 ± 4.4	16	137.5 ± 6.6
Others	62	106.6 ± 2.3	72	121.3 ± 2.3
	$F = 2.06$		$F = 7.73$	
	$P = 0.15$		$P = 0.007$	

Discussion

Based on the seven polymorphic sites studied, the proposed phylogenetic tree consisted of six haplotypes, each derived from its progenitor in the most parsimonious manner to give a simple tree. In the Swedish individuals studied, all these haplotypes were observed unambiguously, and were the most frequent haplotypes seen. However, in addition, another three unpredicted haplotypes were observed unequivocally, and at least another nine, much rarer, haplotypes were also deduced to be present. The finding that these rare, unpredicted haplotypes were significantly more frequent in the patient group than among the controls

may, if confirmed in other studies, be of clinical importance. Our findings are supported by those of Ludwig and McCarthy (Ref. 36, Table III), who investigated apo B haplotypes in eight American kindreds with familial defective apo B100. Using 10 polymorphic markers on 67 chromosomes, these workers found 22 different haplotypes. The only chromosomes that they observed more than twice correspond to the haplotypes that we have numbered 1–6 (inclusive). They also detected two chromosomes that correspond to haplotype 7, and at least one chromosome conforming to each of the rare, unpredicted haplotypes that we designated: 9,10,12,13,16 and 18. Thus, in addition to the haplotypes that we observed unequivocally, this family study also demonstrates the existence of five of the nine haplotypes that we were only able to deduce to be present. There are several possible mechanisms by which these unpredicted haplotypes could have occurred. Recombination events between two different apo B alleles may be the most likely possibility, but it is also possible that some of the polymorphisms considered here are at hypermutable sites and that either the same mutation has occurred twice on two different haplotypes, or that a mutation having occurred once, has subsequently reverted back to its original form.

If recombination is the mechanism by which these unusual haplotypes are generated, it is notable that in all except one of the unpredicted haplotypes seen, the site of recombination would appear to be 5′ to the *Xba*I site. It is notable that more than 99% of the intron sequences within the apo B gene are also positioned 5′ to the *Xba*I site and exon 26 [37]. This data extends our previous findings of a lack of recombination within the 3′ end of the apo B gene (exons 26–29) [18] and implies that apparent recombination events may correlate with the presence of intron sequences, as suggested previously by Gilbert [38]. Taken together these data support a hypothesis that change in the apo B gene is driven more by novel mutation than by recombination.

The control group of individuals in this study demonstrate the frequently-seen association between the *Xba*I RFLP and serum apo B levels. The unpredicted haplotypes 7 and 8, which carry the *Xba*I X+ allele, are associated with a small effect on serum apo B levels in the control group and a significant effect in the patient group. The fact that the effect on apo B levels associated with haplotypes 7 and 8 is greater in the patient sample than in the sample of healthy individuals is unexplained, but if confirmed, is of obvious clinical significance. There is also a similar trend in the association of both the *Xba*I RFLP and haplotypes 7 and 8 with serum LDL cholesterol levels, although these do not reach statistical significance in these Swedish samples. This is consistent with the idea that whilst apolipoprotein levels are largely under genetic control, lipid levels are affected to a greater extent by environmental factors [39].

The analysis of the effect of these haplotypes by quartiles of apo B levels (Fig. 3) suggests that these haplotypes have a mild, but consistent co-dominant effect with more than 40% of individuals carrying these haplotypes having apo B levels in the upper quartile, and 25% having levels above the 90th percentile. This implies that these haplotypes encode common variant forms of apo B protein with a mild functional impact on the individual but a significant effect at the population level.

If haplotypes 7 and 8 do have a functional effect on serum apo B levels, the question is raised as to what is the mechanism of this effect. It could perhaps be due to the presence of a novel mutation, unique to haplotypes 7 and 8, that either alone or in combination with the Ala-591 to Val change, alters the metabolism of the LDL particle or affects the production of apo B protein. Alternatively, if haplotypes 7 and 8 arose by recombination between other haplotypes, the functional effect might be due to the new combinations of existing amino acid polymorphisms in the apo B protein. Robinson and co-workers [40] investigated individuals who have been phenotyped for the Ag loci and reported that although the polymorphisms, considered individually, were not associated with a significant effect on lipid traits, Ag(a/d) in association with either Ag(t/z) or Ag(x/y) was significantly related to plasma IDL mass. It is not clear from this study which haplotypes are associated with the highest IDL levels, however taken together, both studies support the idea that novel combinations of individual amino acid substitutions in apo B may have a greater effect on the function of the protein and thus, on plasma levels of apo B and lipids. This could be the case if the apo B protein has a significant tertiary structure, for which there is growing evidence [41–43]. Studies to ascertain which of these mechanisms is correct are under way.

Acknowledgements

This work was supported by grants from the British Heart Foundation (grant RG16), the Swedish Medical Research Council (grants 08691 and 204), the Swedish Heart-Lung Foundation and the Torsten and Ragnar Söderberg Foundation. H-H.R. was supported by a grant from the Deutsche Forschungsgemeinschaft and A.H. is a cancer investigator of the Swedish Heart-Lung Foundation. We would like to thank Gina Deeley for help in preparing the manuscript.

References

1 Yang, C.Y., Chen, S.H., Gianturco, S.H., Bradley, W.A., Sparrow, J.T., Tanimura, M., Li, W.H., Sparrow, D.A., De Loof, H.,

Rosseneu, M., Proot, L., Brasseur, B., Rugsschaert, J.M., Lee, F.-S., Gu, Z.-W., Grotto, A.M. and Cher, L. (1986) Nature 323, 738–742.
2 Yang, C-Y., Gu, Z-W., Weng, S-A., Kim, T.W., Chen, S-H., Pownall, H.J., Sharp, P.M., Li, S.-W., Li, W.-H., Gotto, A.M. and Chan, L. (1989) Arteriosclerosis 9, 1, 96–108.
3 Hamsten, A., (1988) Acta. Med. Scand. 223, 389–403.
4 Hamsten, A., Iselius, L., Dahlén, G. and De Faire, U. (1986a) Atherosclerosis 60, 199–208.
5 Berg, K. (1987) in Molecular approaches to human polygenic disease (Bock, G. and Collins, G.M., eds.), pp 14–28, Ciba Foundation Symposium; 130, 1st Edn. Wiley & Sons, Chichester, UK.
6 Pairitz, G., Davignon, J., Maillouz, H. and Sing, C.F. (1988) Am. J. Hum. Genet. 43, 311–321.
7 Hasstedt, S.J., Wu, L. and Williams, R.R. (1987) Genet. Epidemiol. 4, 67–76.
8 Innerarity, J.L., Mahley, R.W., Weisgraber, K.H., Bersot, T.P., Krauss, R.M., Vega, G.L., Grundy, S.M., Friedl, W., Davignon, J. and McCarthy, B.J. (1990) J. Lipid Res. 31, 1337–1349.
9 Allison, A.C. and Blumberg, B.S. (1961) Lancet 1, 634–637.
10 Breguet, G., Bütler, R., Bütler-Brunner, E. and Sanchez-Mazas, A. (1990) Am. J. Hum. Genet. 46, 502–517.
11 Ma, Y., Schumaker, V., Bütler, R. and Sparkes, R.S. (1987) Arteriosclerosis 3, 301–305.
12 Ma, Y., Wang, X., Bütler, R. and Schumaker, V.N. (1989) Arteriosclerosis 9, 242–246.
13 Dunning, A.M., Tikkanen, M.J., Ehnholm, C., Bütler, R. and Humphries, S.E. (1988) Hum. Genet. 78, 325–329.
14 Wang, X., Schlapser, P., Ma, Y., Bütler, R., Elovson, J. and Schumaker, V.N. (1988) Arteriosclerosis 8, 429–435.
15 Young, S.G. and Hubl, S.T. (1989) J. Lipid Res. 30, 443–449.
16 Xu, C-F., Nanjee, N., Tikkanen, M.J., Huttunen, J.K., Pietinen, P., Bütler, R., Angelico, F., Del Ben, M., Mazzarello, B., Antonio, R., Miller, N., Humphries, S. and Talmud, P.J. (1989) Hum. Genet. 82, 322–326.
17 Wu, M-J., Bütler, E., Bütler, R. and Schumaker, V.N. (1991) Arteriosclerosis and Thrombosis 11, 379–384.
18 Dunning, A.M., Renges, H-H., Xu, C-F., Peacock, R., Brasseur, R., Laxer, G., Tikkanen, M.J., Bütler, R., Saha, N., Hamsten, A., Rosseneu, M., Talmud, P. and Humphries, S. (1992) Am. J. Hum. Genet. 50, 208–221.
19 Zerba, K.E., Kessling, A.M., Davignon, J. and Sing, C.F. (1991) Genetics 129, 525–533.
20 Boerwinkle, E., Lee, S.S., Bütler, R., Schumaker, V.N. and Chan, L. (1990) Atherosclerosis 81, 225–232.
21 Rapacz, J., Chen, L., Bütler-Brunner, E., Wu, M-J., Hasler-Rapacz, J., Bütler, R. and Schumaker, V. (1991) Proc. Natl. Acad. Sci. USA 88, 1403–1406.
22 Renges, H-H., Peacock, R., Dunning, A.M., Talmud, P. and Humphries, S.E. (1992) Ann. Hum. Genet. 56, 11–33.
23 Berg, K., Powell, L.M., Wallis, S.C., Pease, R., Knott, T.J. and Scott, J. (1986) Proc. Natl. Acad. Sci. USA 83, 7367–7370.
24 Law, A., Wallis, S.C., Powell, L.M., Pease, R.J., Brunt, H., Priestley, L.M., Knott, T.J., Scott, J., Altman, D.G., Miller, G.J., Rajput, J. and Miller, N.E. (1986) Lancet i, 1301–1303.
25 Talmud, P.J., Barni, N., Kessling, A.M., Carlsson, P., Darnfors, C., Bjursell, G. and Galton D. (1987) Atherosclerosis 67, 81–89.
26 Aalto-Setälä, K., Gylling, H., Helve, E., Kovanen, P., Miettinen, T.A., Turtola, H. and Kontula, K. (1989) Hum. Genet. 82, 305–307.
27 Tikkanen, M.J., Xu, C-F., Hamalainen, T., Talmud, P., Sarna, S., Huttunen, J.K. and Pietinen, P. (1989) Clin. Genet. 37, 327–334.
28 Carlsson, P., Darnfors, C., Olofsson, S.O. and Bjursell, G. (1986) Gene 49, 29–51.
29 Hamsten, A., Walldius, G., Dahlén, G., Johansson, B. and De Faire, U. (1986b) Atherosclerosis 59, 223–235.
30 Hamsten, A., Eféndic, S., Walldius, G., Szamosi, A. and De Faire, U. (1987) Am. Heart J. 113, 917–927.
31 Peacock, R., Dunning, A., Hamsten, A., Tornval, P., Humphries, S. and Talmud, P. (1992) Atherosclerosis 92, 151–164.
32 Ludwig, E.H., Blackhart, B.D., Pierotti, V.R., Caiati, L., Fortier, C., Knott, T., Scott, J., Mahley, R.W., Levy-Wilson, B. and McCarthy, B.J. (1989) DNA 6, 4, 363–372.
33 Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T. and Mullis, K.B. (1988) Science 239, 487–491.
34 Chakravarti, A., Buetow, K.H., Antonarakis, S.E., Waber, P.G., Boehm, C.D. and Kazazian, H.H. (1984) Am. J. Hum. Genet. 36, 1239–1258.
35 Xu, C-F., Tikkanen, M.J., Huttunen, J.K., Pietinen, P., Bütler, R., Humphries, S. and Talmud, P. (1990) J. Lipid Res. 31, 1255–1261.
36 Ludwig, E.H. and McCarthy, B.J. (1990) Am. J. Hum. Genet. 47, 712–720.
37 Blackhart, B.D., Ludwig, E.H., Pieratti, V.R., Caiati, L., Onasch, M.A., Wallis, S.C., Powell, L., Pease, R., Knott, J.J., Chu, M.-L., Mahley, R.W., Scott, J., McCarthy, B.J. and Levy-Wilson, B. (1986) J. Biol. Chem. 261, 15364–15367.
38 Gilbert, W. (1985) Science 228, 823–8.
39 Reilly, S.L., Kottke, B.A. and Sing, C. (1990) J. Clin. Epidemiol. 43, 9, 921–940.
40 Robinson, M.T., Bütler, R. and Krauss, R.M. (1991) Clin. Genet. (in press).
41 Walsh, M.T. and Atkinson, D. (1990) J. Lipid Res. 31, 1051–1061.
42 Chatterton, J.E., Phillips, M.L., Curtiss, L.K., Milne, R.W., Marcel, Y.L. and Schumaker, V.N. (1991) J. Biol. Chem. 266, 9, 5955–5962.
43 Koduri, R.K., Austin, R.K., Curtiss, L.K., Young, S.J., Witiztum, J.L. and Smith, R.S. (1991) Arteriosclerosis and Thrombosis 11,5, 1404a.

APOLIPOPROTEIN B: GENETIC VARIANTS PROVIDE INSIGHT INTO STRUCTURE AND FUNCTION

Philippa Talmud, Alison Dunning, Steve Humphries
Dept of Medicine
Rayne Institute
University College Middlesex School of Medicine
London, UK

In order to fulfil its function as a secretory protein, lipid transporter and ligand for the LDL-receptor, apolipoprotein B (apo B) has a number of functional domains; a signal peptide that directs the translocation of the protein across the endoplasmic reticulum to the Golgi apparatus, a region that binds to the LDL-receptor and a domain for lipid binding to maintain the integrity of the lipoprotein particle and stability during lipolysis. Much of our understanding of the structure-function aspects of apo B have come from studies of rare genetic variants. Several genetic disorders of lipoprotein metabolism are directly linked to the apo B gene or associated with the biosynthesis, assembly or secretion of apo B containing lipoproteins.

RARE MUTATIONS

Familial Hypobetalipoproteinaemia (HBL): HBL is an autosomal codominant disorder with heterozygotes having total cholesterol, LDL cholesterol and apo B levels below the 5th centile. Cosegregation studies have shown that the mutation is in the apo B gene (Leppert et al., 1988). All the mutations to date result in truncated proteins and most have been initially identified at the protein level using polyacrylamide gel electrophoresis of plasma or LDL to get an estimate of protein size. In a recent screening carried out on blood donors in St Louis, Missouri, the frequency of HBL was estimated to be less than 0.01% (Wagner et al., 1991). In the homozygous form, patients have trace levels of apo B-containing

although detected in vitro, are probably very unstable and in vivo degrade after secretion. The apo B46 (Young et al, 1989) apo B50 (Hardman et al., 1989) and apo B54.8 (Wagner et al., 1991) are found primarily in the VLDL density range although apo B46 is seen in LDL and HDL ranges and apo B54.8 in LDL density range. The larger apo B species are found in a similar density range as full length apo B100. Thus, the truncated apo B proteins help define the region in the amino-terminal end of the protein important for the association with lipids.

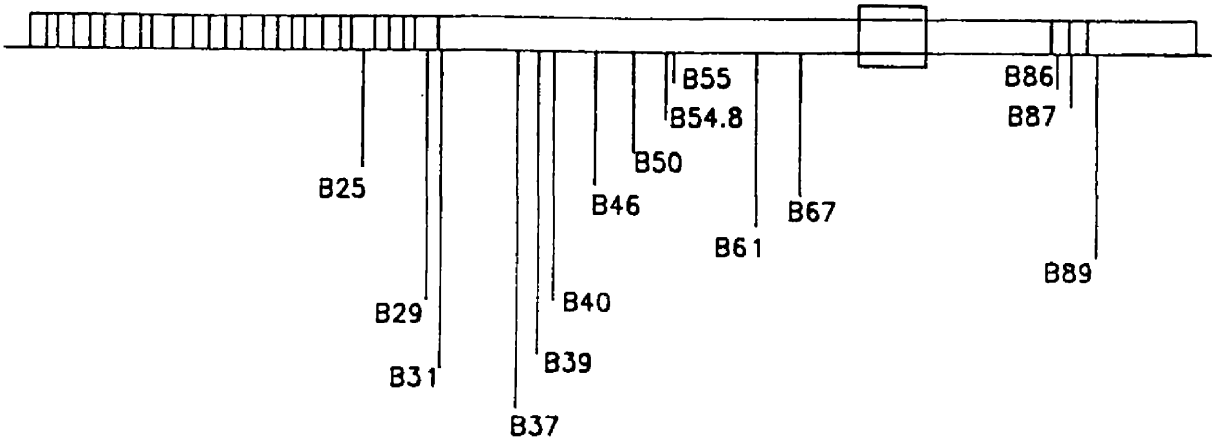


FIGURE 1. Schematic map of the apo B cDNA identifying reported mutations that lead to HBL. Vetical lines on the cDNA map represent the introns. The open box depicts the proposed LDL-receptor binding domain

Definition of additional LDL-R binding domains: The mechanism causing low levels of these truncated forms in the plasma is unclear. Most of the truncated apo B proteins, identified to date, terminate before the putative receptor-binding domain that binds to the LDL-R. Thus an increase in catabolism is unlikely to be an explanation for the low concentrations of apo B and some other mechanism/s must explain this. Instability of the mutant lipoproteins may contribute to their low levels. The larger truncated proteins apo B87 and apo B89, which do encode the primary LDL-R binding domain provide evidence of a secondary binding domain on apo B for the LDL-R. In the case of the apo B89, degradation and binding were increased on cultured fibroblasts when compared to apo B100 (Krul et al., 1989) and an enhanced clearance of apo B89 in LDL-turnover studies performed in rabbits, suggests that

lipoproteins often associated with neuropathies and retinopathies, the result of fat-soluble vitamin deficiency. The extent of the severity appears to be related to the size of the truncated apo B protein and whether it can associate with lipid.

These truncated apo B species help define the functional domains of apo B, namely the lipid-binding and receptor-binding regions. There is now accumulated evidence about the lipid binding capabilities of these truncated proteins from the density range in which they are identified, that provides insight into their lipid binding potential.
More than 15 truncated apo B mutations causing HBL have now been characterised at the molecular level. A map of the apo B gene showing the position of the mutations that lead to truncated proteins appear in Figure 1. Except in the case of the apo B25, the result of a deletion of the entire exon 21, all the truncated forms reported to date are due to C→T transitions or base deletions. In addition all the reported mutations have been unique to the kindreds they have been identified in and no two unrelated families have the same mutation.

Definition of a lipid-binding domain: Where the mutation creates a protein greatly reduced in size, for example the truncated apo B species apo B25 (Huang et al., 1989) and apo B29 (Collins et al., 1988) no plasma apo B was detected. It has been suggested that these small truncated species lack a domain necessary for lipoprotein association and stability. The amino acid residues in the amino terminus of the protein are hydrophilic and would not bind well to lipid. The truncated species apo B31 (Young et al., 1990) and apo B37 (Young et al, 1987) are associated with small amounts of triglyceride-rich lipoproteins and occur primarily in the HDL density range; this reflects the small amount of associated lipid and not an HDL-type particle.

In vitro expression studies have further clarified the lipid binding domain. Expression of cDNA constructs in either transfected rat hepatoma cell lines or HepG2 cells have clarified that constructs smaller than apo B28 associate to a minor extent with lipid but are found primarily in the infranatant (Yao et al, 1991; Graham et al, 1989). Thus there appears to be a critical size which the protein must exceed in order to associate with lipid. In the region of apo B31 and apo B37 there is an increased amount of hydrophobic amino acids which can bind to lipid. However in view of the fact that naturally occurring apo B25 and apo B29 are not detected in the plasma it would seem that truncated apo B proteins smaller than apo B31,

apo B89 has an increased affinity for the LDL receptor (Parhofer et al., 1990). The absence of the carboxy-terminal 500 amino acids in apo B89, which are hydrophobic and firmly embedded in lipid, may result in the release of the constraint put on the primary receptor binding domain or make more available a secondary binding region. The functional domains defined by the truncated proteins are shown in Figure 2

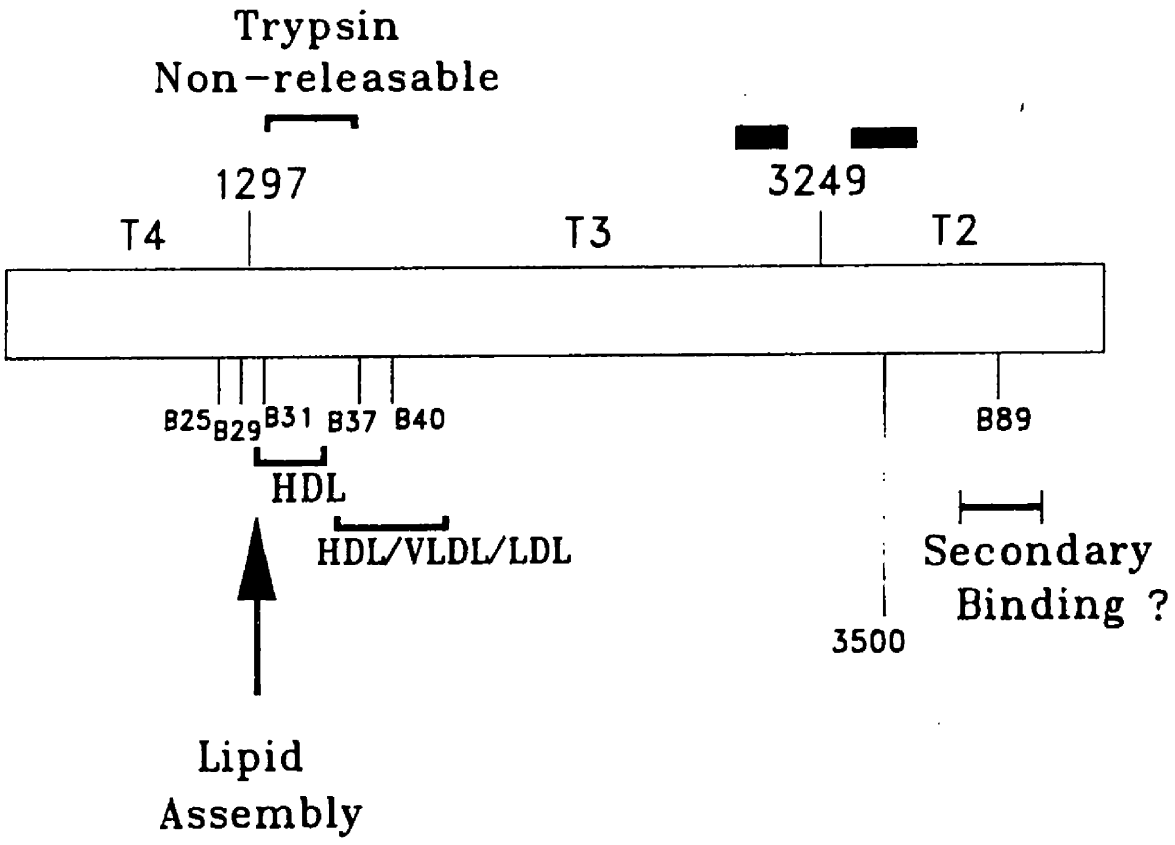


Figure 2. Schematic map of the apo B protein showing the thrombolytic fragments. The filled blocks represent the putative receptor binding domain. Lipid binding and receptor binding domains defined further by the truncated proteins are shown.

Why are plasma apo B levels so low in heterozygous HBL patients? It is still unclear , however, some heterozygous HBL patients have LDL levels below the 5th centile and not at the 50th centile, despite the presence of one normal apo B100 allele. Some insight comes from LDL-turnover studies in a patient with apo B55 (Talmud et al., 1992) which revealed a reduced LDL synthetic rate (5.5mg/dl/day) compared to the normal range of 10-15mg/dl/day (Converse et al., 1990). Apo B synthesis is constitutive in Hep G2 cells and the control of apo B output is thought to be either due to an increase in the mRNA translational efficiency or the decreased degradation of intracellular apo B in the presence of oleate

(Pullinger et al,. 1989). Thus it is possible that in the hepatocytes of the heterozygous HBL patients, while mRNAs coding for both the truncated and normal apo B proteins are translated, the truncated species interfere with the assembly of apo B100 into normal lipoproteins, leading to an overall reduction in plasma apo B-containing lipoproteins.

Familial Defective apo B100 (FDB): It is well known that defects in the LDL-R that destroy function lead to FH. Similarly, defects in the apo B gene that would reduced binding to the LDL-R could result in raised cholesterol levels. The disorder Familial Defective apo B100 (FDB) has been proposed for such mutations. A number of laboratories have carried out systematic searches, using molecular biology techniques, to identify patients who have mutations in the apo B gene that would alter receptor binding. To date there is only one form of FDB and that is due to a mutation that substitutes codon Arg_{3500} by Gln, hence called the apo B3500 mutation. Originally observed in a patient who showed reduced clearance of autologous LDL, when compared to clearance of LDL from a normal donor the mutation was identified as a G→A substitution altering codon CGG to CAG (Soria et al., 1989). This single amino acid change reduces binding of the LDL containing apo B-Gln_{3500} resulting in the accumulation of such LDL while LDL-Arg_{3500} is cleared with normal efficiency. The frequency of FDB has been estimated to be 1/500-1/700. Although residue 3500 is 120 amino acids to the COOH-terminal side of the initially defined receptor binding domain, the Arg→Gln substitution clearly affects the interaction with the receptor. It was observed that monoclonal antibody MB47 binds with a greater affinity to the defective LDL-Gln_{3500} suggesting that the Arg_{3500}→Gln substitution has a large effect on the conformation of that region of the protein (Weisgraber et al., 1988). Furthermore Lund-Katz et al (1989), using ^{13}C NMR studies have shown that the six lysine residues within the region of amino acid 3500 have altered pKs in the presence of Gln_{3500}. Lysine residues are known to be involved in the binding of apo B100 to the LDL-R and the substitution of Arg_{3500}→Gln by altering the pKs of the lysines, alters the microenvironment of the receptor binding domain.

Thus mutations in patients with HBL or FDB help define both the lipid-binding and receptor-binding domain of apo B and provide evidence that sequence changes outside the primary binding domain may affect the affinity of LDL for the receptor. Furthermore the molecular defects leading to HBL illustrate the genetic heterogeneity that underlies this syndrome.

COMMON POLYMORPHISMS

Association studies with common polymorphisms of the apo B gene have identified those polymorphisms which show a consistant association with differences in lipid levels. The association studies themselves do not provide information as to whether the polymorphisms are functional or whether the effects seen are due to linkage disequlibrium with functional sequence changes elsewhere in the gene. We have been particularly interested in identifying the functional sequence changes that underlie these effects.

XbaI polymorphism: Several Restriction Fragment Length Polymorphisms (RFLPs) of the apo B gene have been identified and using these RFLPs, population studies have shown that variation at a polymorphic XbaI site within the apo B gene is associated with differences in serum cholesterol levels, with the X+ allele (presence of the XbaI cutting site) being associated with raised levels (reviewed by Humphries, 1988). Although this association has not been seen in every study, it has been observed in many different laboratories, using samples from different countries, and the consistent nature of this finding makes it highly unlikely that the observation is due to chance alone. The size of the effect associated with this RFLP is modest explaining 3 - 8% of the sample variance in cholesterol levels, and is thus of the same order of magnitude as is seen with variation at the apo E gene (Davignon et al., 1988). The base change that creates the XbaI site does not alter an amino acid, as it is in the third (wobble) position of a codon for threonine 2488, and so cannot be the direct cause of the effects seen. We thus postulate a second sequence variation, elsewhere in the gene and in linkage disequilibrium with the XbaI RFLP, that leads to altered serum lipid levels. This sequence difference could occur in the promoter region of the apo B gene where it may alter transcription of the gene and thus affect production of the protein from the intestine or liver. Alternatively, as with apo E2, it may change an amino acid in the protein and alter the affinity of LDL for its receptor. Evidence in favour of the latter mechanism has come from two in vivo studies that have shown that compared with the X- allele, the X+ allele is associated with a reduced fractional catabolic rate (FCR) of autologous LDL (Demant et al., 1988; Houlston et al., 1988). These data thus predict that LDL from individuals with the genotype X-X- would bind to the LDL-receptor with increased affinity compared to LDL from individuals with the genotype X+X+. Support for this prediction has been obtained using a competative binding and internalisation assay of radiolabelled LDL

to normal fibroblasts (Series et al., 1989), where "X+" LDL competed poorly compared to "X-" LDL, as expected from the turnover data. This is compatible with the hypothesis that a common amino acid change in the apo B protein, may affect the affinity for the LDL receptor.

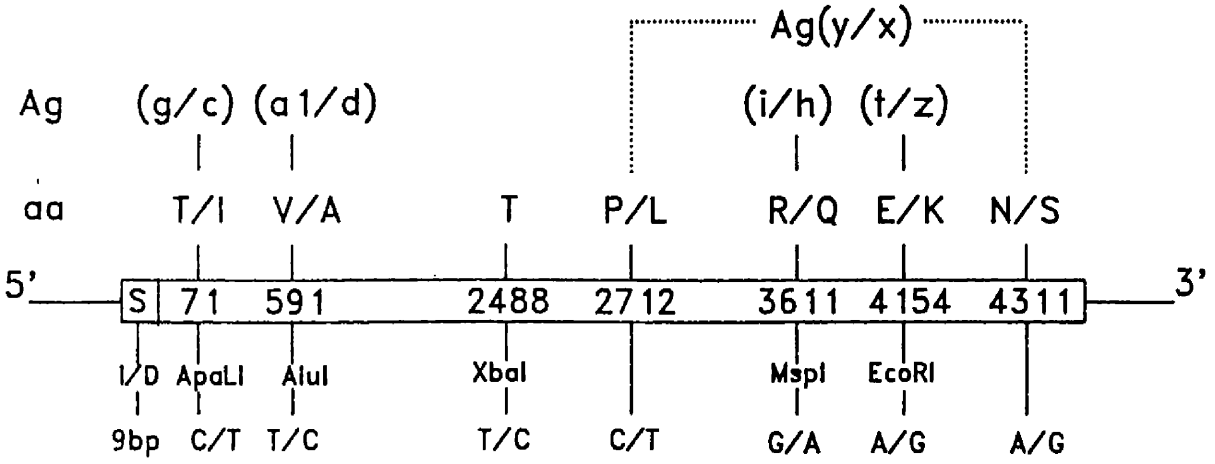


Figure 3. Map of the apo B gene showing the relative position of the identified apo B Ag polymorphisms, the signal peptide polymorphism, and the XbaI RFLP variable site.

Ag polymorphisms: Several common protein polymorphisms of apo B are known (Fig 3). These were first identified in 1961 as Antigenic determinants (Ag) using antibodies found in multiple transfused individuals (Allison and Blumberg, 1961). Five pairs of epitopes were identified by these studies, with each pair representing alleles at closely linked loci (Breuget et al., 1990). With the cloning and sequencing of the apo B gene, the molecular basis of four of these polymorphisms has been deduced from association studies. The c/g and a1/d, are located in the N-terminal region of the protein which is trypsin-accessable and to be the result of single base changes in the gene that cause respectively the substitution of Ile for Thr (Ma et al., 1989) and Val for Ala (Wang et al., 1990) amino acid. The c/g epitopes have been associated with differences in plasma cholesterol levels (both LDL-C and HDL-C) in a large study of children from Finland (Tikkanen et al., 1988), but this association was not observed in an earlier smaller study (Young et al., 1987). No association between the a1/d epitopes and lipid levels have been reported. The Ag epitope pairs h/i and t/z occur in the COOH-terminal part of the protein, in peptides that are not released by digestion of LDL

The Pro-Asn / Leu-Ser polymorphism shows strong allelic association with the XbaI polymorphism with the Leu-Ser (Ag x) being found exclusively on chromosomes lacking the XbaI site. In a sample of 90 healthy, Swedish individuals the Leu_{2712}/Ser_{4311} allele is associated with reduced serum levels of LDL cholesterol and apo B and with raised levels of High Density Lipoprotein, thus confirming the finding of Berg et al., (1976). However these differences are of smaller effect than those associated with the XbaI Restriction Fragment Length Polymorphism (RFLP) of the apo B gene in this sample (Dunning et al., 1992). It thus appears that although the two amino acid changes are associated with some effect on LDL-function, they do not explain all the variation seen with the XbaI RFLP and further functional genetic polymorphism at this locus remain to be discovered.

Signal peptide polymorphism: Another example of a common polymorphism with a possible functional role is the insertion/deletion polymorphism of the apo B signal peptide which has been shown to be associated with differences in triglyceride levels (Xu et al, 1990) and baseline atherosclerosis score (Peacock et al, 1992). Whether the presence or absence of the three amino acids seen in the insertion and deletion alleles respectively have a direct effect and affect the rate of translocation across the endoplasmic reticulum or whether the association is to an indirect effect due to linkage disequlibrium with a second sequence change. We will test the hypothesis by expressing these signal peptide alleles, linked to a yeast secretory protein in a yeast expression system.

Thus both the rare mutations and the common polymorphisms of the apo B gene help define the those regions of the apo B protein with functional properties. The consistency of the association between the polymorphism at the apo B locus and differences in plasma levels of apoprotein and lipoprotein suggest that in addition, when the functional sequence changes are identified these genetic markers will be useful in assessing an individuals risk of developing hyperlipidaemia and thus CAD.

ACKNOWLEDGEMENTS

Support for this work was provided by the British Heart Foundation (grant RG16), and the Charing Cross Sunley Research Trust. The authors would like to thank Ms Gina Deeley for help in preparing the manuscript.

by trypsin. The h/i polymorphism at residue 3611 is of interest because, like the apo B 3500 mutation (Innerarity, 1990), it is the result of a change from Arg to Gln. This amino acid region is likely to be located on the surface of the LDL-particle since it alters an epitope recognised by antibodies and is close to the region of apo B that interacts with the LDL-receptor. However, unlike apo B 3500, Arg_{3611}-Gln does not appear to alter the affinity of the particle for the LDL-receptor, and has not been consistently associated with a significant effect on lipid levels (Xu et al., 1989; Rajput-Williams et al., 1988). The t/z polymorphism is unique in that it substitutes a positively charged amino acid Lys for a negatively charged Glu (Ma et al., 1987). Associations have been noted between this polymorphism and differences in VLDL-triglyceride levels and plasma lipid levels although this association has not been confirmed in other studies, (see Humphries, 1988). Until recently, the position of the remaining Ag epitope pair x/y had not been determined. This polymorphism has been associated with differences in plasma cholesterol and triglyceride levels (Berg et al., 1976), but the effect is small, and only reached statistical significance after combining data from a large number of individuals (n = 1087) from different studies. On average, individuals carrying one or more "x" allele had cholesterol levels roughly 3% lower and triglyceride levels 16% lower than those with only the "y" allele.

To identify the sequence creating the x/y epitope a search was made for sequence differences between a group of 10 individuals (Dunning et al., 1991) using PCR amplification of different exons of the gene followed by chemical mismatch cleavage. An A to G base change was identified in exon 29 of the apoB gene that results in the substitution of serine for asparagine at residue 4311 of mature apo B100 (Dunning et al., 1992). In a recent publication, Huang et al (1990) have reported a C to T base change in exon 26 that causes the substitution of leucine for proline at residue 2712 of apo B. We have found complete allelic association between the alleles at both these sites and Ag (x/y) in a sample of 118 Finnish individuals (37). This implies that either one, or both of these substitutions combined, could be the molecular basis of the Ag(x/y) antigenic determinants - the allele encoding Leu_{2712} plus Ser_{4311} representing the Ag(x) epitope and that encoding Pro_{2712} plus Asn_{4311} the Ag(y). It is also possible that the epitope is created by an additional sequence change yet undetected in the gene that is in complete association with the changes at residues 2712 and 4310. One approach to distinguish these three possibilities would be by expressing in HepG2 cells, constructs of the apo B gene containing the different amino acid substitutions.

REFERENCES

Allison AC, Blumberg BS (1961) An isoprecipitation reaction distinguishing human serum protein types. Lancet 1 1:634-637

Berg K, Hames C, Dahlen G, Frick MH, Krishan I (1976) Genetic variation in serum low density lipoproteins and lipid levels in man. Proc Natl Acad Sci USA 73:937-940

Breguet G, Butler R, Butler-Brenner E, Sanchez-Mazas A (1990) A worldwide population study of the Ag-system haplotypes, a genetic polymorphism of human low-density lipoprotein. Am J Hum Genet 46:502-517

Collins DR, Knott TJ, Pease RJ, Powell LM, Wallis SC, Robertson S, Pullinger CR, Milne RW, Marcel YL, Humphries SE, Talmud PJ, Lloyd JK, Miller NE, Muller D, Scott J (1988) Truncated variants of apolipoprotein B cause hypobetalipoproteinemia. Nucl Acid Res 16:8361

Converse CA, Keegan WA, Huq L, Series J, Caslake M, McLachlan J, Packard CJ, Shepherd J (1990) Further Epidemiological Studies in Lipid Metabolism in Retinitis Pigmentosa, *in:* Inherited and Enviromentally Induced Retinal Degeneration, M.M.Lavail, J.G.Hollyfield and R.E.Anderson (Eds), Eds, Alan M Liss, New York

Davignon J, Gregg RE, Sing CF (1988) Apolipoprotein E polymorphism and atherosclerosis. Arteriosclerosis 8:1-21

Demant T, Houlston RS, Caslake MJ et al (1988) Catabolic rate of low density lipoprotein is influenced by variation in the apolipoprotein B gene. J Clin Invest 82:797-802

Dunning AM, Houlston R, Frostegard J, Revill J, Nilsson J, Hamsten A, Talmud P, Humphries S (1991) Genetic evidence that the putative receptor binding domain of apolipoprotein B (residues 3130 to 3630) is not the only region of the protein involved in interaction with the low density lipoprotein receptor. Biochem Biophys Acta 1096:231-237

Dunning AM, Tikkanen MJ, Enholm C, Butler R, Humphries SE (1988) Relationships between DNA and protein polymorphisms of apolipoprotein B. Hum Genet 78:325-329

Dunning AM, Renges H-H, Xu C-F, Peacock R, Brasseur R, Laxer G, Tikkanen MJ, Butler R, Saha N, Hamsten A, Rossenue M, Talmud P and Humphries SE (1991) Complete Allelic association between two amino acid substitutions in apolipoprotein B and the x/y antigen group polymorphism. Am J Hum Genet (In Press).

Graham LD, Pullinger CR, Jones TC, Knott TJ, Scott J (1989) Expression and secretion of apolipoprotein (apo)-B constructs by HepG2 cells. Circulation 80:II:1854

Hardman DA, Pullinger CR, kane JP, Malloy MJ (1989) Molecular defect in normotriglyceridaemic abetalipoproteinaemia, Circulation, 80:(II) 466

Houlston RS, Turner PR, Revill J, Lewis B, Humphries SE (1988) The fractional catabolic rate of low density lipoprotein in normal individuals is influenced by variation in the apolipoprotein B gene. Atherosclerosis 71:81-85

Huang LS, Gavish D, Breslow JL (1990) Sequence polymorphism in the human apo B gene

at position 8344. Nucleic Acids Res 18:5922.

Huang L-S, Ripps ME, Korman SH, Deckelbaum RJ, Breslow, JL (1989) Hypobetalipoproteinemia due to an apolipoprotein B gene exon 21 deletion derived by Alu-Alu recombination. J Biol Chem 264:11394

Humphries SE (1988) DNA polymorphisms of the apolipoprotein genes -their use in the investigation of the genetic component of hyperlipidaemia and atherosclerosis. Atherosclerosis 72:89-108.

Krul ES, Kinoshita M, Talmud P, Humphries SE (1989) Turner, S., Goldberg, A. C., Cook, K., Boerwinkle, E., and Schonfeld, G., Two distinct truncated apolipoprotein B species in a kindred with hypobetalipoproteinemia. Arteriosclerosis 9:856

Leppert M, Breslow JL, Wu L, Hasstedt S, O'Connell P, Lathrop M, Williams RR, White R, Lalouel J-M (1988) Inference of a molecular defect of apolipoprotein B in hypobetalipoproteinemia by linkage analysis in a large kindred. Clin Invest 8266:847

Lund-Katz S, Innerarity TL, Curtiss LK, Krauss RM, Phillips MC (1989) C-NMR evidence that substitution of Gln for Arg 3500 in familial defective Apo B-100 disrupts the conformation of the receptor binding domain. Arteriosclerosis 9:715a

Ma Y, Wang X, Butler R, Schumaker VN (1989) Bsp 12861 restriction fragment length polymorphism detects Ag(c/g) locus of human apolipoprotein B in all 17 persons studied Arteriosclerosis 9:242-246.

Ma Y, Schumaker V, Butler R, Sparkes RS (1987) Two DNA restriction fragment length polymorphism (RFLPs) associated with Ag (t/z) and Ag (c/g) antigenic sites of human apolipoprotein B. Arteriosclerosis 3:301-305.

Parhofer KG, Daugherty A, Kinoshita M, chonfeld G (1990) Enhanced clearance from plasma of low density lipoproteins containing a truncated apolipoprotein, Apo B-89. J Lipid Res 31:2001

Peacock R, Dunning A, Hamsten A, Tornvall P, Humphries S, Talmud P (1992) Apolipoprotein B gene polymorphisms, lipoproteins and coronary atherosclerosis: A study of young myocardial infarction survivors and healthy population-based individuals. Atherosclerosis 92:151-164.

Pullinger CR, North JD, Teng BB, Rifici VA, Ronhild de Brito AE, Scott J (1989) The apolipoprotein B gene is constitutively expressed in HepG2 cells: regulation of secretion by oleic acid, albumin, and insulin, and measurement of the mRNA half-life. J Lipid Res 30:1065

Rajput-Williams J, Knott TJ, Wallis SC et al (1988) Variation of apolipoprotein B-gene is associated with obesity, high blood cholesterol levels , and increased risk of coronary heart disease. Lancet 1442-1446

Series J, Cameron I, Caslake M, Gaffney D, Packard CJ, Shepherd J (1989) The XbaI polymorphism of the apolipoprotein B gene influences the degradation of low density lipoprotein in vivo. Biochim Biophys Acta 1003:183-188

Soria LF, Ludwig EH, Clarke HR, Vega GL, Grundy SM, McCarthy BJ (1989) Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. Proc Natl Acad Sci USA 86:587

Talmud PJ, Converse C, Krul E, Haq L, McIlwaine GG, Series JJ, Boyd P, Schonfeld G, Dunning A, Humphries SE (1992) A novel truncated apolipoprotein B (apo b55) in a patient with hypobetalipopriteinaemia and autosomal dominant retinitis pigmentosa. Clinical Genetics (Submitted)

Talmud P, King Underwood L, Krul E, Schonfeld G, Humphries S (1989) The molecular basis of truncated forms of apolipoprotein B in a kindred with compound heterozygous hypobetalipoproteinemia. J Lipid Res 30:1773-1779

Tikkanen MJ, Viikari J, Kakerblom HK, Pesonen E (1988) Apolipoprotein B polymorphism and altered apolipoprotein B and low density lipoprotein cholesterol concentrations in Finnish children. Brit Med J 296:169-170

Wagner RD, Krul ES, Tang J, Parhofer KG, Garlock K, Talmud P, Schonfeld G (1991) Apo B 54.8, a truncated apolipoprotein found primarily in VLDL is associated with a nonsense mutation in the apo B gene and hypobetalipoproteinemia. J Lipid Res 32:1001

Wang X, Schlapfer P, Ma Y, Butler R, Elovson J, Schumaker VN (1990) Apolipoprotein B: The Ag($_1$/d) immunogenetic polymorphism coincides with a T-to-C substitution at nucleotide 1981, creating an alu 1 restriction site. Arteriosclerosis 8:429-435

Weisgraber KH, Innerarity TL, Newhouse YM, Young SG, Arnold KS, Krauss RM, Vega G L (1988) Familial defective apolipoprotein B-100: enhanced binding of monoclonal antibody MB47 to abnormal low density lipoproteins. Proc Natl Acad Sci USA 85:9758

Xu C-F, Nanjee N, Tikkanen MJ et al (1989) Apolipoprotein B amino acid 3611 substitution from arginine to glutamine creates the Ag (h/i) epitope: the polymorphism is not associated with differences in serum cholesterol and apolipoprotein B levels. Hum Genet 82:322-326

Xu C-F, Tikkanen MJ, Huttunen JK, Pietinen P, Butler R, Humphries SE, Talmud P (1990) Apolipoprotein B signal peptide insertion/deletion polymorphism is associated with AG epitopes and involved in the determination of serum triglyceride levels J.Lipid Res. 31: 1255-1261.

Yao Z, Blackhart BD, Linton MF, Taylor SM, Young SG, McCarthy BJ (1991) Expression of carboxyl-terminally truncated forms of human apolipoprotein B in rat hepatoma cells : evidence that the length of apolipoprotein B has a major effect on the buoyant density of the secreted lipoproteins. J Biol Chem 266:3300

Young SG, Hubl ST, Smith RS, Snyder SM, Terdiman JF, (1990) Familial hypobetalipoproteinemia caused by a mutation in the apolipoprotein B gene that results in a truncated specied of apolipoprotein B (B-31). A unique mutation that helps to define the portion of the apolipoprotein B molecule required for the formation of byoyant,triglyceride-rich lipoproteins. J Clin Invest 85:933

Young SG, Bertics SJ, Scott TM et al (1987) Apolipoprotein B allotypes MB19$_1$ and MB$_2$ in subjects with coronary artery disease and hypercholesterolemia. Arteriosclerosis 7:61-65

Young SG, Bertics SJ, Artiss LK, Witztum JL (1987) Characterization of an abnormal species of apolipoprotein B, apolipoprotein B-37, associated with familial hypobetalipoproteinemia. J Clin Invest 79:1831

Young SG, Hubl ST, Chappel DA, Smith RS, Claiborne F, Snyder SM Terdiman JF, (1989) Familial hypobetalipoproteinemia associated with a mutant species of apolipoprotein B (B-46). N Engl J Med 32:1604