

UNIVERSITY OF CAPE TOWN



**The molecular and cellular defect underlying
autosomal recessive hypercholesterolemia (ARH) in
the first kindred identified in South Africa.**

Faculty of Health Sciences

Department of Medicine

Student Name: Aron Abera
Student Number: ABRARO001
Project Supervisors: Profs HE Henderson and AD Marais

Submission for MSc degree: March 2005

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

TABLE OF CONTENTS

<u>CONTENT</u>	<u>PAGE</u>
ACKNOWLEDGMENTS	I
LIST OF ABBREVIATIONS	II
PROJECT SYNOPSIS	V
CHAPTER 1: INTRODUCTION	I
1.1 Cholesterol	2
1.2 Lipoproteins	2
1.3 Familial Hypercholesterolemia	8
1.4 Autosomal Recessive Hypercholesterolemia	12
1.7 Project Aim	25
CHAPTER 2: MATERIALS AND METHODS	27
2.1 Patients	28
2.2 Cellular Investigations	31
2.3 Molecular Studies	38
CHAPTER 3: RESULTS	51
3.1 Patient I	52
3.2 Patient II	67
CHAPTER 4: DISCUSSION & CONCLUSION	72
APPENDIX A	78
APPENDIX B	81
REFERENCES	82

ACKNOWLEDGEMENTS

- I would like to express my sincere thanks to my supervisors, Professors Howard Henderson and Dave Marais for their guidance throughout the duration of the project and for financial support.
- To Felicity Leisegang, Sheena Jones and Ingrid Baumgarten for their assistance in the laboratory.
- To Prof. Linton Traub (University of Pittsburgh School of Medicine, USA) for his kindly provision of anti-ARH antibody.
- To Prof. J Kastelein (University of Amsterdam, Holland) for generously giving us access to DNA samples from the Iranian ARH family.
- To Prof. Helen Hobbs (University of Texas Southwestern medical center, USA) and Prof. Anne Soutar (MRC Clinical Sciences Center, UK) for helpful discussions.
- To my girlfriend Kim for being understanding throughout the work.
- Last, but not least, I would like to extend a special thanks to the staff members of the Division of Lipidology and Chemical Pathology (Medical School, UCT) for their friendly assistance throughout the project.

LIST OF ABBREVIATIONS

A	adenine
A _β PP	amyloid β-precursor protein
apo -	apolipoprotein
APS	ammonium persulphate
ARH	autosomal recessive hypercholesterolemia
bp	base pair
BSA	bovine serum albumin
C	cytosine
cDNA	complementary DNA
D	aspartic acid
Dab	disabled protein
DEPC	diethylpyrocarbonate
dH ₂ O	distilled water
Dif	1,1'-dioctadecyl-3,3,3'-tetramethylindocarbocyanine perchlorate
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphates
DMEM	Dulbecco's Modification of Eagle's Minimum Essential Medium
DMSO	dimethylsulphoxide
E	glutamic acid
EBV	Epstein-Barr virus
EDTA	ethylenediamine tetraacetic acid
ER	endoplasmic reticulum
F	phenylalanine
FCS	fetal calf serum
FDB	Familial defective apoB-100
FH	familial hypercholesterolemia
g	acceleration due to gravity
G	guanine
HDL	high density lipoprotein

HIFCS	heat-inactivated fetal calf serum
HL	hepatic lipase
HSPG	heparan sulfate proteoglycans
I	isoleucine
IDL	intermediate density lipoprotein
KBr	potassium bromide
KDa	kilodalton
kb	kilobase
L	leucine
LDL	low density lipoprotein
LDLr	low density lipoprotein receptor
LPDP	lipoprotein deficient plasma
LPDS	lipoprotein deficient serum
LPL	lipoprotein lipase
LRP	LDL receptor-related protein
mRNA	messenger ribonucleic acid
mg	milligram
ml	milliliter
N	asparagine
NPC	Nieman-Pick type C
P	proline
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PCSK9	proprotein convertase subtilisin/kexin type 9
PHA	phytohaemagglutinin
PMSF	phenylmethylsulfonyl fluoride
PTB	phosphotyrosine binding
rpm	revolutions per minute
RT	reverse transcriptase
SDS	sodium dodecyl sulphate

SNP	single nucleotide polymorphism
SSCP	single strand conformation polymorphism
T	thymine
TAE	tris-acetate EDTA
TBE	tris-boric acid EDTA
TEMED	N, N, N', N'-tetramethylethylenediamine
Tris	tris (hydroxymethyl) amino methane
V	valine
VLDL	very low density lipoprotein
X	stop codon
Y	tyrosine

University of Cape Town

PROJECT SYNOPSIS

Inefficient clearance of the circulating cholesterol enriched lipoprotein, low density lipoprotein (LDL), leads to elevated plasma levels and hypercholesterolemia, cholesterol ester accumulation in the vascular bed and premature atherosclerosis. LDL is cleared, mainly by the liver, through interaction with LDL receptors (LDLr) located on the surface of hepatocytes. Monogenic defects in the LDL uptake pathway occur commonly in South Africans, particularly in the Afrikaner community where inheritance is typically autosomal dominant, arising predominantly from abnormal structure and thus function of the LDLr. Defects in LDLr binding domain of apolipoproteinB-100 (apoB-100) are rarely encountered and are known as Familial defective apoB-100 (FDB).

Several critical proteins are active in the LDL uptake pathway and their deficiencies are now being shown to underlie the rare autosomal recessive forms of hypercholesterolemia (ARH). One of these proteins is the LDLr adaptor protein known as ARH, which is presumed to facilitate interaction of the cytoplasmic tail of the LDLr with the internal protein matrix required for the receptor internalisation.

The aim of this project was firstly to confirm the provisional clinical diagnosis of ARH in two patients who presented with the clinical and laboratory features of homozygous FH, but with atypical family histories and secondly, to characterize any mutation found in the ARH gene (*ARH*) in terms of PCR detection, haplotype determination, family and general population carrier screening.

The first patient investigated, Patient I was originally seen at the Lipid Clinic of Johannesburg General Hospital in 1993, at the age of 27 years. Blood was taken for DNA isolation with a skin biopsy being taken for fibroblast culture. LDL uptake kinetics were

determined in cultured skin fibroblasts and found to be normal, with results of 99.4% of the control fibroblast cultures. The coding sequence of *ARH* was determined by cDNA sequencing, with the cDNA being generated by RT-PCR. These sequencing studies revealed a single guanine base (G) insertion in a run of 7 G nucleotides in the first exon of *ARH*. This insertion changes the reading frame and introduces a premature stop codon at amino acid position 33 of the ARH protein. This mutation thus significantly truncates ARH. The establishment of a PCR based assay for the mutation allowed for carrier detection in 400 black South African newborns. None of these samples tested positive for the insertion mutation. Two single nucleotide polymorphisms (SNPs) were also found in the patient, one in exon 6 and one in exon 7. The SNP in exon 6 at position 604 of the cDNA (CCC→TCC; proline for serine) occurs in black South Africans at a frequency of 0.37 and 0.63 for the 'C' and 'T' alleles, respectively. The second SNP in exon 7 at position 654 (ACG→ACA), is silent as it does not alter the amino acid sequence. This SNP is also novel, as it does not appear in any SNP database. The alleles of this SNP, 'G' and 'A' have a frequency of 0.82 and 0.18 respectively, in black South Africans. Both SNPs were used for haplotype determination to allow for a comparison of the allelic markers associated with the South African and previously described identical Iranian 'G' insertion mutations. These data revealed these two mutations to have different genetic backgrounds, most likely being indicative of different mutational events.

The second South African patient investigated, patient II, is a 5-year-old boy of mixed ancestry, who presented with markedly elevated plasma LDL cholesterol and clinical features of homozygous FH. However, his parents presented with lipid profiles only slightly elevated and not typical of heterozygous FH and he was thus regarded as a good

candidate for an ARH defect. As this family is of mixed ancestry it was expedient to search for the 'G' insertion mutation detected in our black patient. This mutation was excluded by use of the in-house PCR assay developed in this study, diminishing slightly the likelihood of ARH deficiency.

This likelihood was further reduced when haplotype analysis revealed patient II to be heterozygous for two SNP markers for *ARH*. Patients with autosomal recessive disorders are most often homozygous for intragenic markers due to consanguinity or founder gene inheritance. Being heterozygous was felt to diminish further the likelihood of ARH deficiency.

The final investigative procedure carried out in this kindred, was a Western blot to look at ARH synthesis in transformed lymphocytes. A good full length ARH signal was obtained from the proband cell extract, strongly suggesting that ARH deficiency was an unlikely possibility in this family. However, it must be remembered that all we have shown is the presence of ARH protein; Western blotting makes little comment on the functional activity of the detected proteins.

In conclusion, this study characterizes the genetic defect in the first kindred affected by ARH in South Africa. The *ARH* defect has been shown to be a frame-shift mutation in exon-1, which severely truncates ARH, indicating a complete absence of ARH function in the proband. Deficiency of ARH was also shown by Western blotting where fibroblast protein extracts failed to yield a signal when probed with antibodies to ARH. The PCR assay set up to detect this insertion mutation can easily be used to screen other patients with suspected ARH deficiency or to determine the prevalence of this mutation in selected cohorts.

ARH deficiency has been shown to be unlikely in the second kindred and proband examined (patient II), although a primary or secondary defect in ARH function remains a possibility and will require further study.

University of Cape Town

CHAPTER 1:

INTRODUCTION

1.1 CHOLESTEROL	2
1.2 LIPOPROTEINS	2
1.2.1 Classes of lipoproteins	4
1.2.2 Lipoprotein Metabolism	4
1.2.3 LDL-receptor and Lipoprotein Metabolism	6
1.3 FAMILIAL HYPERCHOLESTEROLEMIA	8
1.3.1 Clinical Features	9
1.3.2 Genetics Basis of FH	11
1.4 AUTOSOMAL RECESSIVE HYPERCHOLESTEROLEMIA (ARH)	12
1.4.1 Clinical Features	14
1.4.2 Genetics Basis of ARH	14
1.4.3 LDL-receptor Adaptor (ARH) Protein	15
1.4.4 ARH and the Endocytic Machinery	19
1.4.5 Mutations in the <i>ARH</i> Gene	21
1.4.6 <i>ARH</i> Mutation Demographics	24
1.4.7 <i>ARH</i> mRNA Stability	24
1.4.8 <i>ARH</i> Polymorphisms	25
1.4.9 Treatment	25
1.5 PROJECT AIM	25

1.1 CHOLESTEROL

“Cholesterol is a rigid hydrophobic molecule that gives structural integrity to plasma membranes of vertebrate cells” (Rader et al. 2003). It was first isolated from gallstones in 1784 and since then has exerted great attraction for scientists from diverse areas of study. Thirteen Nobel prizes have been awarded to various scientists who have devoted their careers to this molecule. Cholesterol has a major function in membranes of animal cells, where it modulates fluidity and keeps the barrier between cells and its environment. It is also used as a raw material for the production of steroid hormones and bile acids (Brown & Goldstein 1986).

Free cholesterol, in deficiency or excess, is toxic to cells. Intracellular levels are thus maintained within critical limits by a number of processes. These are:

1. **de novo synthesis:** endogenous cholesterol is synthesized through several intermediates from acetyl CoA as the precursor.
2. **Cholesterol uptake (Influx):** exogenous cholesterol is taken up as a component of various lipoproteins by a method discussed in section 1.2.3.
3. **Efflux:** cells protect themselves from cholesterol overloading by export into the extracellular compartment.
4. **Storage:** excess cholesterol is stored in the form of cholesteryl ester droplets.
5. **Metabolism:** Conversion to steroid hormones, bile acids or vitamin D.

1.2 LIPOPROTEINS

Cholesterol is not soluble in plasma making transport within the circulation difficult; the body however, solves this problem by esterifying the sterol with long-chain fatty acids

and by carrying these esters within the hydrophobic cores of lipoproteins (Brown et al. 1986). Lipoproteins are molecules that are used to transport lipids, mainly cholesterol and triglyceride, and fat soluble vitamins and are classified according to their densities and protein composition. These particles are spherical and have a central core of nonpolar lipids (primarily triglycerides and cholesteryl esters) and a surface monolayer of polar lipids (primarily phospholipids) and apoproteins. The ratio and the composition of these molecules determine their size and density. The density of a lipoprotein particle is inversely related to its size, reflecting the relative amounts of low density, nonpolar core lipid and high density, surface protein present (Scriver et al. 1995).



Fig. 1.1: Structure of lipoprotein with cholesteryl ester in the core and phospholipids in the surface monolayer. (Reproduced from www.people.vcu.edu/~urdesai/ahlp.htm)

1.2.1 Classes of Lipoproteins

The lipoproteins of largest size contain mostly triglycerides in their cores. These are the chylomicrons, secreted from enterocytes, in which the apolipoprotein B (apoB) is primarily or exclusively apoB-48, and the very low density lipoprotein (VLDL), secreted by hepatocytes, which contain apoB-100. Much smaller sized are the low density lipoprotein (LDL) and high density lipoprotein (HDL₂ and HDL₃), which contain mainly cholesterol esters in their cores. The mature form of these lipoproteins are not secreted directly from cells but rather are produced by metabolic process within the blood plasma. LDLs are produced as end products or remnants of the metabolism of VLDL. The sixth class, intermediate density lipoprotein (IDL), which contains appreciable amounts of both triglycerides and cholesterol esters in the core, is produced during the conversion of VLDL to LDL (Scriver et al. 1995).

1.2.2 Lipoprotein Metabolism

Lipoprotein metabolism is a very broad subject but this review concerns itself with the LDL receptor (LDLr) and is thus limited to the metabolism of apoB containing lipoproteins, chylomicrons and VLDLs.

Chylomicron

Dietary cholesterol and triglycerides are packaged with apolipoproteins in the endoplasmic reticulum (ER) of enterocytes of the small intestine, secreted into lymphatic system as chylomicrons. The protein components of chylomicrons include apoB-48 and other apoproteins. As they circulate, the core triglycerides are hydrolyzed by lipoprotein lipase (LPL) resulting the formation of chylomicron remnants. The remnants bind to

hepatic lipoprotein receptors including the LDLr and LDL receptor-related protein (LRP), and are then rapidly removed by the liver (Rader et al. 2003).

Dietary cholesterol has three possible fates once it reaches the liver.

1. It can be esterified and stored as cholesteryl esters in hepatocytes.
2. Packaged into VLDL particles and secreted in to the plasma or
3. Converted into bile acids and secreted to the bile.

Cholesterol is transported from peripheral cells to the liver by HDL and is recycled to LDL and VLDL by cholesterol-ester transfer protein (CETP) or is taken up by the liver through interaction with HDL (Scriver et al. 1995).

VLDL

VLDLs are synthesized and assembled in hepatocytes and contain mainly apoB-100 and small amounts of apoE. The initial phase of the metabolism is the same as the chylomicrons. VLDL remnants interact with LDLr and LRP on hepatocytes via apoE and are taken up by the liver through LDLr-mediated endocytosis. Some of the remnants undergo further lipolysis by hepatic lipase (HL) and mature to LDL, the major cholesterol carrying particle in the blood. An estimated 70% of circulating LDL will be internalized and degraded primarily by the liver using a sole ligand apoB-100 (Mahley et al. 1996). The terminal catabolism of lipoproteins that contain apoB happens by receptor-mediated endocytosis. The best-studied lipoprotein receptor is LDLr and it mediates the internalisation of particles containing apoB-100. This receptor interacts with ARH (autosomal recessive hypercholesterolemia), a putative adaptor protein that is involved in the mechanics of the endocytic process.

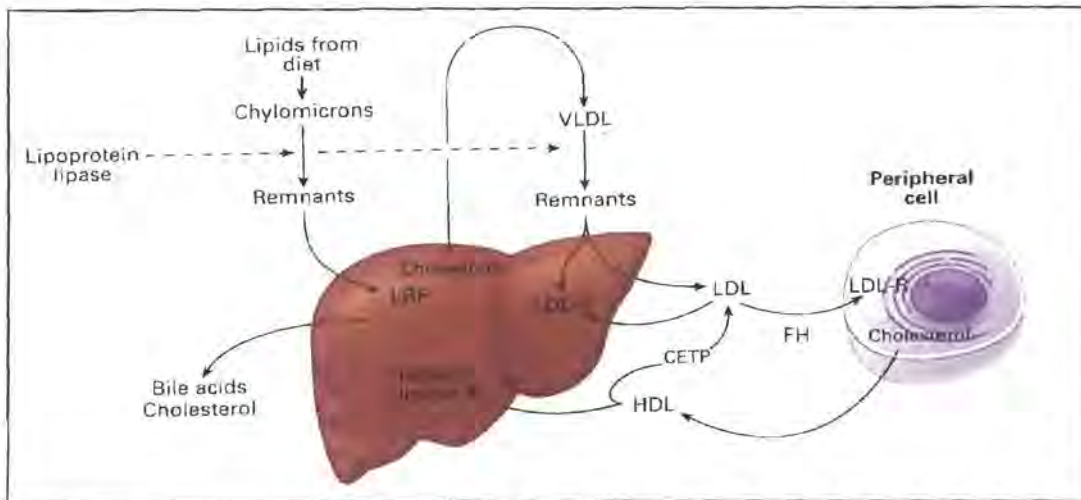


Fig. 1.2: A brief overview of lipoprotein metabolism in humans. (Reproduced from Broeders et al. 1999).

1.2.3 LDL receptor and Lipoprotein Metabolism

The human LDL receptor (LDLr) is a transmembrane cell-surface glycoprotein and it was first discovered by Brown and Goldstein in 1970s. It plays an important role in internalisation of LDL by binding two protein ligands: 1) ApoB-100 (a sole protein of LDL) and 2) ApoE, that is found primarily in VLDL and chylomicron remnants as well as IDL and a subclass of HDL. LDL-receptors on interacting with their lipoprotein ligands, gather on the cell surface in coated pits. These pits are specialized regions of the plasma membrane that are lined on the cytoplasmic surface by a protein called clathrin. Clathrin is a molecule in the coated pits that is responsible for retaining the LDLr during receptor-mediated endocytosis. The cytoplasmic tail of the LDLr forms direct and indirect bonds with the heavy chain of clathrin (Kibbey et al. 1998). Within a few minutes of their formation, the coated pits invaginate to form coated endocytic vesicles. Multiple endocytic vesicles then connect to create larger structures of irregular contours called endosomes. Thereafter, the LDL dissociates from the receptor and is delivered to the lysosomes for degradation by acid hydrolytic enzymes. The receptor in turn recycles

to the surface for repeat interaction with LDL. The apoprotein and cholesteryl esters of LDL are hydrolyzed by protease and lysosomal acid lipase respectively while unesterified cholesterol will be transported by Neiman-Pick type C (NPC) 1 and 2 transport proteins to the lysosomal membrane and enters the cellular compartment (Scriver et al. 1995; Frolov et al. 2003).

The five structural domains of the LDLr were largely determined by the Goldstein and Brown group and are shown in figure 1.3 (Brown & Goldstein 1986). This review will focus on the cytoplasmic domain of the receptor and its interaction with some adaptor proteins. A conserved hexapeptide amino acid sequence in the cytoplasmic tail of LDLr (FDNPVY) has been shown to be necessary for receptor clustering in coated pits and for endocytosis. This hexapeptide signal sequence, lies within the first 22 amino acids of the cytoplasmic tail and directs LDL receptors to the coated pits. The validity of this sequence has been shown by site-directed mutagenesis studies and by naturally occurring LDLr gene mutations, which have impaired receptor function. Residue 807, which is the terminal tyrosine, has a critical requirement for an aromatic side chain; amino acid substitutions other than tryptophane or phenylalanine impair receptor function significantly (Chen et al. 1989).

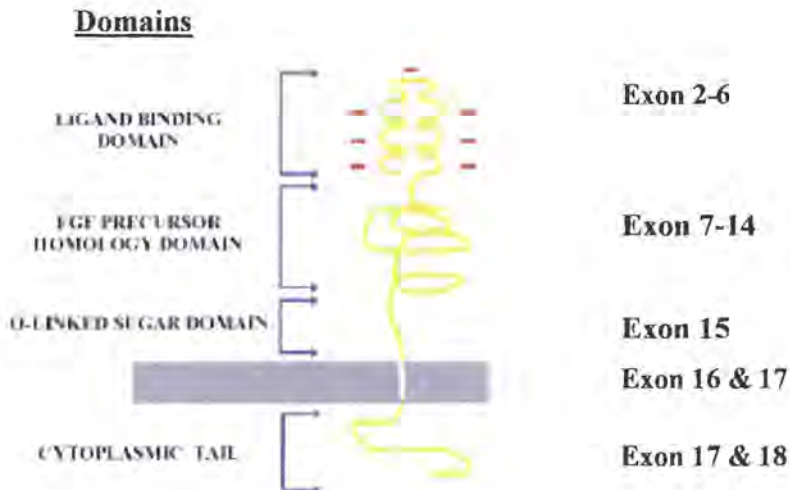


Fig. 1.3: The structural domains of the LDL-receptor and their corresponding exons. (Reproduced from Brown & Goldstein 1986)

The concentration of plasma LDL-cholesterol (LDLc) will thus rise if there is any defect in the interaction of LDL with its receptor or in the clustering of occupied receptors in coated pits and their internalisation. Plasma levels of LDLc are directly associated with the occurrence of coronary events and cardiovascular deaths. The variation of plasma LDLc levels in the normal population comes from polygenic genetic variation, showing the interactive effect of multiple gene sequence variants with environmental factors in any given individual. Patients with very high LDLc levels have a monogenic form of hypercholesterolemia that is associated with accumulation of cholesterol in tissues, skin and tendons called xanthomata, and atheroma in the vascular system, which causes atherosclerosis (Rader et al. 2003).

1.3 FAMILIAL HYPERCHOLESTEROLEMIA

Familial hypercholesterolemia (FH) is the most severe form of monogenic hypercholesterolemia and it was the first genetic disease of cholesterol metabolism to be clinically and molecularly distinguished. The correlation of atheromas in arteries and

xanthomata in tendons was repeatedly described before 1900, with the monogenic autosomal dominant nature of the disorder being shown by the extensive observations of Khachadurian in a Lebanese family (Khachadurian 1964). In 1970s Brown and Goldstein discovered the cell-surface LDLr and its association with FH. They also showed binding of LDL to LDLr served for two purposes these being, a negative feedback mechanism for cholesterol biosynthesis by suppression of HMG-CoA reductase and an enhancement of endocytosis of LDL for further degradation in lysosomes (Brown & Goldstein 1986).

1.3.1 Clinical Features

The diagnosis of this disorder is based on family history of hypercholesterolemia and premature coronary artery disease but clinical features vary according to the differing ages of affected individuals, genetic variation and other factors. A gene dosage effect is evident in this disorder with homozygotes being more severely affected than heterozygotes.

FH is clinically characterised by:

1. The deposition of LDL-derived cholesterol in skin and tendon (xanthomata), affecting predominantly the Achilles tendon, hands and elbows, and arteries (atheroma).
2. Premature coronary artery disease
3. A high plasma concentration of LDL, the main cholesterol-transporting lipoprotein.
4. An autosomal dominant pattern of inheritance (Scriver et al. 1995).

Severe hypercholesterolemia is evident at birth in homozygotes with yellow-orange cutaneous xanthomas appearing as early as two years of age. The development of tendon

xanthomas, arcus cornealis and atherosclerosis are predictable in childhood; death from myocardial infarction usually occurs before the age of 30 (Scriver et al. 1995; Rader et al. 2003).

Table 1.1: Plasma lipid levels in Caucasian patients with Familial Hypercholesterolemia (Reproduced from Scriver et al. 1995; Marais et al. 2004; Kotze et al. 1993)

General Population Data						
Genotype	Age, yrs	Plasma Cholesterol, mmol/L				Plasma
		Total	VLDL	LDL	HDL	Triglyceride, mmol/L
Normal	1-19	4.5 ± 0.7	0.3 ± 0.2	2.8 ± 0.6	1.4 ± 0.3	0.7 ± 0.3
Heterozygotes	1-19	7.7 ± 1.6	0.4 ± 0.3	6.2 ± 1.5	1.1 ± 0.3	0.9 ± 0.6
Homozygotes	1-19	17.4 ± 4.4	0.5 ± 0.2	16 ± 4	0.9 ± 0.3	1.1 ± 0.6
Normal	≥20	5 ± 0.9	0.4 ± 0.3	3.2 ± 0.8	1.4 ± 0.4	0.9 ± 0.3
Heterozygotes	≥20	9.4 ± 2	0.7 ± 0.4	7.6 ± 2	1.1 ± 0.3	1.7 ± 0.8
South African Data						
Genotype	Age, yrs	Plasma Cholesterol, mmol/L				Plasma
		Total	VLDL	LDL	HDL	Triglyceride, mmol/L
Normal	27-55	5.7 ± 1.3	-	3.8 ± 1.2	1.4 ± 0.4	1.8 ± 1.5
Heterozygotes	27-53	9.4 ± 1.5	-	7.7 ± 1.4	1.1 ± 0.3	1.3 ± 0.6
Homozygotes	2-43	18.1 ± 3.7	-	16.2 ± 3.3	0.8 ± 0.3	1.3 ± 0.9

In heterozygotes, the earliest manifestation is a high concentration of cholesterol in plasma, which remains the only clinical/laboratory finding throughout the first decade. Arcus cornealis and tendon xanthomas start to appear at the end of the second decade and each are present in at least half of the heterozygotes by the third decade. The fourth decade is associated with the clinical manifestation of coronary heart disease and, by the time of the death almost 80% of the heterozygotes have xanthomas (Scriver et al. 1995).

1.3.2 Genetics Basis of FH

FH is transmitted as an autosomal dominant trait with a background prevalence reported to be 1/500 in Caucasians (heterozygotes). Much higher prevalences are however, seen in founder populations in three areas of the world. In Lebanese and French Canadians the carrier frequency is 1/171 and 1/270 respectively, while in South Africa the Afrikaner and Ashkenazi Jew communities have frequencies of 1/100 and 1/67 respectively (Scriver et al. 1995; Rader et al. 2003).

There are two genetic causes that are associated with FH. The first and most common one is a disorder that involves a mutation in the LDLr gene. FH patients with these mutations are divided in two major groups according to their LDLr activity in cultured fibroblasts; patients with less than 2% of normal LDLr activity are called receptor negative while those with 2-25% of normal LDLr activity are receptor defective (Scriver et al. 1995). The LDLr gene is located on the short arm of chromosome 19 (p13.1-p13.3), spans 45kb and has 18 exons and 17 introns; to date about 900 mutations have been described in this gene. These mutations fall into four classes (Brown & Goldstein 1986) that can be distinguished by structural criteria those being,

Class 1 Mutations: No receptors synthesized. This class is the most common one and genes with this class of mutations produce either no LDLr or only trace amounts.

Class 2 Mutations: Receptor synthesized, but transported slowly from ER to golgi. This is the second most common class of mutations and these genes produce LDL receptors that are synthesized as precursors with molecular weights that vary from 100,000 to 135,000 (120,000 being the normal precursor size). The receptors that are synthesised

from these mutant genes do not appear on the surface of the cell; rather they seem to stay in the ER until they are eventually degraded.

Class 3 Mutations: Receptors are processed and reach the cell surface, but fail to bind LDL normally. Mature proteins produced from this class of mutation have a normal apparent molecular weight and reach the cell surface. However, they have a markedly reduced ability to bind with LDL. These mutations may involve amino acid substitutions, deletions or duplications in the cysteine rich LDL binding domain.

Class 4 Mutations: receptors reach cell surface and bind LDL, but fail to cluster in coated pits. Studies of this class of mutation have shown the importance of coated pits in receptor-mediated endocytosis. Mutations in this class are predominantly found in the cytoplasmic tail of LDLr and impair internalisation of the receptors.

The second category of defect underlying FH is Familial defective apoB-100 (FDB), which is associated with a missense mutation (Arg3500Gln) in the LDLr-binding domain of apoB. Other less frequent mutations in apoB-100 are also associated with the disease. FDB has a 1/1,000 prevalence ratio in central Europe and is not common in other populations. The clinical features of both dominant disorders are similar but are milder for FDB. Recently proprotein convertase subtilisin/kexin type 9 serine protease (PCSK9) has been described to underlie FH but very little is known about the mechanism of the protein (Rader et al. 2003).

1.4 AUTOSOMAL RECESSIVE HYPERCHOLESTEROLEMIA (ARH)

This genetically distinct form of hypercholesterolemia came into focus in 1973 when Khachadurian described a Lebanese family in which all the children had clinical features similar to homozygous FH and elevated plasma LDLc. The difference however, was that

these children had a slightly reduced or normal LDLr activity in cultured fibroblasts (Khachadurian et al. 1973). Several families subsequently accumulated over the years with FH clinical phenotypes but with autosomal recessive rather than autosomal dominant patterns of inheritance (Zuliani et al. 1995). The first evidence of a defect in LDLr function came from a cellular study on 2 unrelated pairs of siblings, both products of consanguineous unions that showed a defect in internalisation of LDL in Epstein-Barr virus (EBV)-transformed lymphocytes despite the expression of normal LDLr mRNA and protein (Norman et al. 1999; Eden et al. 2002). In cells from these patients LDLr was synthesized, transported to the cell surface and was able to bind LDL. However, the receptor-LDL complexes accumulate on the cell surface rather than being internalised. Similar uptake studies on monocyte macrophages from these patients revealed impaired LDL uptake. These findings were in contrast to the fibroblasts studies where receptor function was only partially impaired (Figure 1.4). Involvement of the LDLr and apoB genes was subsequently discounted by linkage studies, which showed an absence of association with the hypercholesterolemia phenotype.

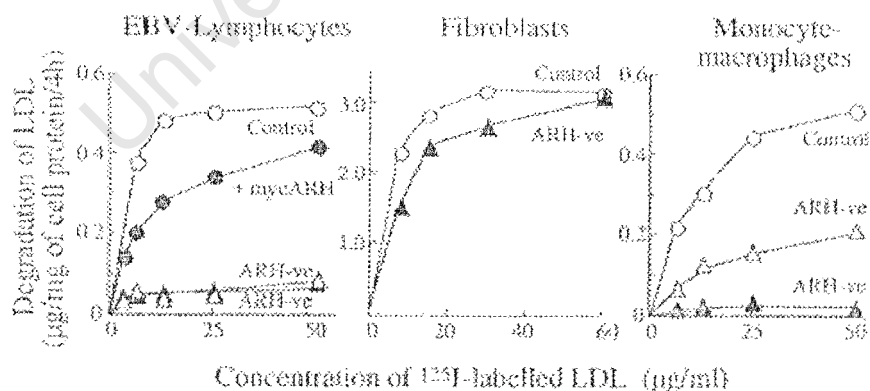


Fig. 1.4: "Degradation of LDL by different cell types. Degradation of ¹²⁵I-labelled LDL by EBV-transformed lymphocytes, skin fibroblasts, and monocyte-derived macrophages from the two ARH (ARH-ve) siblings (filled and open triangles) and from unrelated normolipidemic control subjects (open circles)" (Adapted from Eden et al. 2002).

1.4.1 Clinical Features

The clinical features of ARH are similar to homozygous FH but less severe and of later presentation. Affected individuals develop bulky skin and tendon xanthoma and premature coronary artery disease, which can be fatal if left untreated (Garcia et al. 2001). ARH patients typically manifest with plasma LDLc levels between 9.6-14.3 mmol/L, which is lower than that of homozygous FH 12-20 mmol/L and total plasma cholesterol of 13 –17 mmol/L (Zuliani et al. 1999).

1.4.2 Genetic Basis of ARH

The underlying genetic defect of ARH was discovered recently through studies done in four Sardinian and Lebanese families containing individuals who had a typical clinical and biochemical presentation of ARH. Linkage analysis in these families revealed a significant linkage to a 5.7-cM interval on chromosome 1p35. After refining the linked region to ~1-cM, all the coding sequences of 13 genes in this interval were screened for sequence variation by PCR and by single strand conformation polymorphism (SSCP) analysis. Two abnormally migrating bands were identified in the same gene from both families. Sequencing revealed the nature of the mutations underlying the SSCP band shifts and led to the identification of *ARH* (GenBank accession # AL117654). The gene spans 25kb, has 9 exons and 8 introns. The mRNA encodes for a 308 amino acid protein that has a phosphotyrosine binding (PTB)-domain that binds to the cytoplasmic tail of the LDLr (Garcia et al. 2001; Eden et al. 2001).

It is clear that other gene defects underlie the ARH phenotype but their identity is still unknown. For example, subsequent to the initial linkage studies, linkage analysis in Druze families with ARH (Southern Syria), have shown evidence of linkage to both the

ARH locus (chromosome 1) and to a locus on chromosome 13q22-32. Mutation analysis in the ARH gene revealed a homozygous splice site mutation in intron 1 (Al-Kateb et al. 2002). The nature of the gene defect on chromosome 13 is unknown but it may be significant that another linkage study (Knoblauch et al. 2000) has identified an association between a locus on chromosome 13q and a cholesterol-lowering gene in a FH family with unexpectedly lower concentrations of LDL. In other Sardinian families, there was a strong evidence of linkage to chromosome 15q25-q26 (Ciccarese et al. 2000).

1.4.3 LDL-receptor Adaptor (ARH) Protein

The ARH protein is touted as a putative LDLr adaptor protein but its function is not well understood. It shows high sequence identity with mouse, frog and zebrafish ARH proteins. From this similarity, four conserved segments have been identified that are necessary for effective LDL internalisation (Mishra et al. 2002).

1. The first 44 amino acids of the N-terminal domain
2. A phosphotyrosine binding domain (PTB; residues 45-175)
3. A clathrin box (pentapeptide LLDLE, residues 212-216)
4. Adaptor-2 binding domain (residue 250-280)

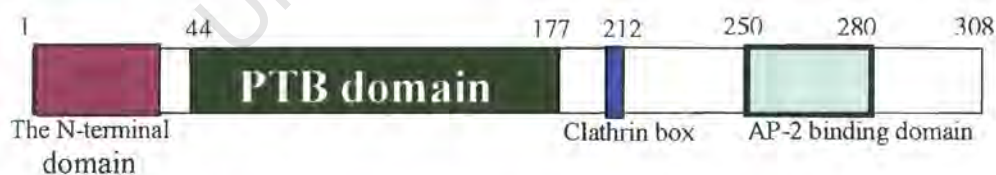


Fig. 1.5: Schematic representation of ARH showing the overall domain structures of the protein; The N-terminal domain (1-44), PTB domain (44-177), clathrin box (212-216) and AP-2 binding domain (250-280). These domains will be discussed below. (Reproduced from Mishra et al. 2002).

1. The N-terminal 44 Amino Acids

This N-terminal portion of the protein consists of a highly conserved ~40 amino acids sequence of unknown function (He et al. 2002). The amino acid sequence alignment is shown in figure 1.6.

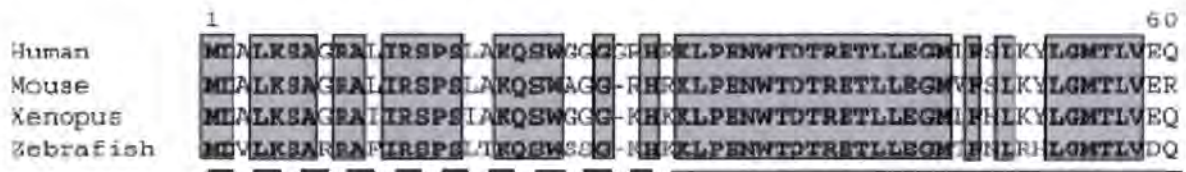


Fig. 1.6: The N-terminal amino acid sequence alignment of the human, mouse, frog and zebrafish ARH proteins. Identical amino acids in all species are boxed (He et al. 2002).

2. The PTB Domain

PTB domains are present in different adaptor proteins that participate in intracellular signaling and transport. The domain of this protein shares some similarity to the endocytic protein Numb (46% similarity) and disabled-2 (Dab2) (46% similarity). “Numb and Dab-2 are endocytic proteins that appear to expand the sorting repertoire of clathrin-coated regions at the cell surface” (Mishra et al. 2002). The presence of a PTB domain suggests that ARH binds with LDLr because the cytoplasmic tail of the receptor carries the hexapeptide sequence FDNPVY, where the consensus PTB hexapeptide is FXNPXY. Although the tyrosine residue in this sequence is normally phosphorylated it would appear that the LDLr interaction with the PTB domain of ARH is not dependent on a phosphorylated tyrosine. A similar situation exists for the ARH homologues, Dab-1 and -2, which also bind to the non-phosphorylated hexapeptide sequence (Chen et al. 1989).

The understanding of the site of interaction between ARH and LDLr was based on a model depicting the PTB domain of X11 interacting with a peptide from amyloid precursor protein (APP) (QNGYENPTYKFFEQ).

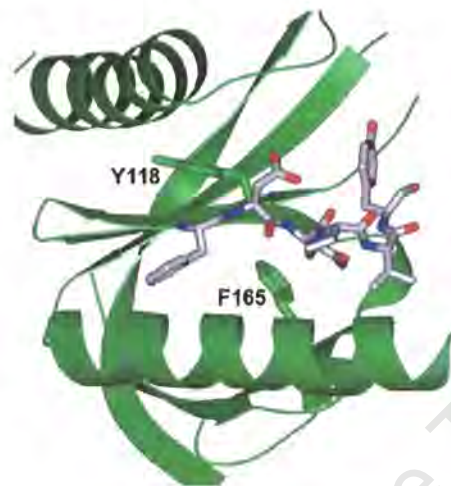


Fig. 1.7: A model structure of ARH PTB domain (green ribbon) in association with the FDNPVY sequence of LDL-receptor. The model is based on the structure of X11 and amyloid precursor protein (APP) (Taken from He et al. 2002)

The derived model for ARH reveals that the two aromatic residues in the PTB domain, Tyr-118 and Phe-165 create a hydrophobic channel that cradles the peptide back bone of FDNPVY sequence in the APP peptide as shown in the figure 1.7. Mutagenesis and transfection studies in which Phe-165 has been substituted by valine or alanine show that while ARH synthesis is not impaired, the protein fails to interact with the LDLr. In addition, conversion of Tyr-118 to cysteine or alanine markedly reduced the interaction of the receptor with ARH (He et al. 2002).

The terminal tyrosine of the NPVY sequence in LDLr tail has a critical requirement for an aromatic side chain; amino acid substitutions other than tryptophane or phenylalanine impair receptor function significantly (Chen et al. 1989).

These data show that the interaction between ARH and the cytoplasmic tail of LDLr requires an aromatic residue at position 807, and that phosphorylation of this very residue is not required for internalisation process (He et al. 2002).

3. The Clathrin Box

ARH also carries a pentapeptide sequence (LLDLE) known as a type-I clathrin box. These are found in many clathrin binding proteins where the consensus sequence is L (L/V/I) (D/E/N) (L/F) (D/E) (Mishra et al. 2002). The last residue of the clathrin box, glutamic acid, interacts with Arg-64 and Lys-96 in the terminal domain of clathrin (He et al. 2002). The importance of the LLDLE sequence has been shown by mutagenesis studies where substitution of the first two leucines or the two acidic residues with alanines, gave attenuation but not complete abolition of clathrin binding (Mishra et al. 2002). These data indicate that ARH has a functional clathrin box sequence, which is likely to be a factor in the ARH dependent retention of LDL/LDLr complexes in clathrin-coated pits (He et al. 2002).

4. Adaptor-2 Binding Domain (Residues 250-280)

The binding of ARH to clathrin is further enhanced by the presence of a C-terminal binding site on ARH for adaptor protein-2 (AP-2), which has clathrin binding properties in it's own right. This has been shown by truncation mutagenesis of ARH, where AP-2 assisted binding of ARH to clathrin is abolished by removal of the terminal 45 amino acids from ARH (He et al. 2002).

AP-2 has several domains that mediate the binding process with other proteins. The β 2-adaptin domain of AP-2 has a large truck domain and a smaller appendage domain that are connected by a residue hinge region as shown in figure 1.8. The hinge region harbors

a clathrin box that mediates the interaction between AP-2 and clathrin. The β 2-subunit appendage binds directly to the AP-2 binding domain of ARH (Figure 1.8). In this domain there are two conserved arginine residues (Arg-261 and Arg-266 of ARH) and a mutation in the first arginine didn't affect the binding activity while the change of Arg-226 to alanine completely abolished the binding process (He et al. 2002).

1.4.4 ARH and the Endocytic Machinery

All these findings prove that the conserved motifs of ARH can mediate protein-protein interactions related to those observed in other endocytic adaptor proteins and is consistent with both genetic and cell biology studies that show the importance of ARH in LDLr endocytosis (He et al. 2002). ARH may enhance the endocytosis of LDLr and LDLr-LDL complexes in coated pits by stabilizing the interaction between the receptor and the structural component of the pit. All ARH domains discussed above make the endocytic machinery complete by their interaction with proteins such as LDLr, AP-2 and clathrin, collectively enabling the accumulation of LDLr in coated pits and successful internalisation of bound LDL as shown in figure 1.8. Generally ARH acts as a molecular escort gathering molecules that have FXNPXY sequences (e.g. LDL-receptor) and attaching them on the clathrin lattice of the coated pits and enhancing the efficiency of packaging into the carrier vesicle (Michaely et al. 2004).

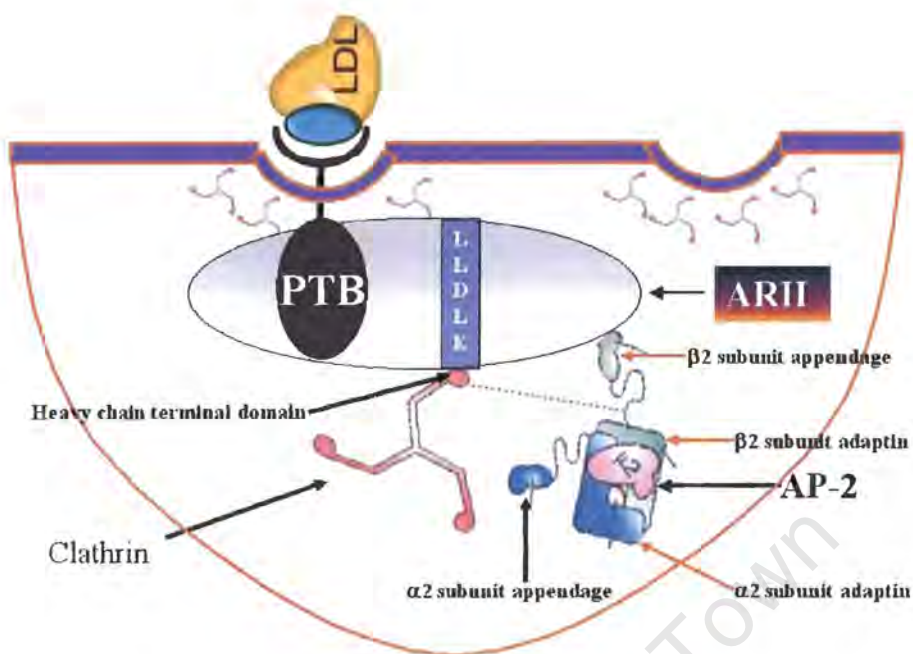


Fig.1.8: A schematic diagram of the major players, and their interactions, in the uptake of LDL through coated pits on the cell surface. Shown are the interactions of ARH with LDLr, Clathrin and AP-2, which enable the successful internalisation of LDL (A composite diagram with some components taken from Rader et al.2003; Soutar et al. 2003).

Studies on fibroblasts from ARH patients have clearly shown that LDL-uptake and degradation is normal or only marginally impaired (Garcia et al. 2001). This most likely indicates that the mechanism of LDL/LDL_r uptake differs from that of the uptake process in hepatocytes and transformed lymphocytes. To account for this, either an adaptor protein is not required in fibroblasts or the role of ARH is compensated for by another sorting protein. ARH fibroblasts have been shown to express the adaptor protein, Dab-2, in high concentration, suggesting that it may take on the ARH role in fibroblasts (Mishra et al. 2002).

The defect in LDL_r function in ARH seems to be not only receptor-specific, but also tissue-specific (Garcia et al. 2001). Liver cells express a much higher amount of LDL_r

than any other cells, yet the expression of ARH is almost constant in all tissues even in those, which have only very low LDLr expression. This may indicate that ARH may have a function other than LDL uptake. Evidence for this may come from two ARH patients of Japanese origin, who were reported to have a fatty liver, which has never been reported as a clinical feature either in FH or ARH. These patients were diagnosed with fatty liver at the age of 30, in the absence of common causes of fatty liver. This finding may thus be related to an unknown function of ARH, although the possibility of an inconsequential association cannot be discounted (Harada-Shiba et al. 2003).

Further, it has recently been shown that ARH interacts with and regulates the cell surface level of Alzheimer's Amyloid β -precursor protein ($A\beta$ PP). This occurs either through regulation of ARH in internalisation of $A\beta$ PP or through transportation of the $A\beta$ PP to the cell membrane. Presumably, ARH deficiency impairs the normal endocytosis of LDLr family members thereby reducing the cholesterol uptake and disrupting the normal regulation of $A\beta$ PP processing (Noviello et al. 2003)

1.4.5 Mutations in *ARH*

To date 13 mutations have been identified in the ARH gene. These have been found in patients from Sardinia, Italy, Lebanon, Iran, USA, Japan, Syria and Mexico, with the bulk of the subjects coming from Sardinia as shown in table 1.2 (Soutar et al. 2003; Canizales-Quinteros et al. 2005).

Table 1.2: Known mutations in *ARH* (updated and adapted from Soutar et al. 2003)

Mutation ¹	No. Of families	Exon	Ethnic origin	Predicted protein
Ins G ₆₅	1	1	Iranian	Fs, G24 + 7X
Del G ₆₅	1	1	American	Fs, G23 + 31X
del GG 86 _{5, 66}	1	1	Pakistani	Fs, G24 + 31X ²
G ₈₆ A (ARH2)	9	1	Sardinian/S. Italian	W22X
	5		Sardinian	
G ₈₆ A + Ins 432 A (ARH3)	1	1	Sardinian	W22X
Ins 2.6Kb	1	IVS1	Lebanese	Null
G ₁ C	1	2 (IVS1)	Syrian Druze	Not known
C ₄₀₆ T	3	4	Turkish/Lebanese	Q136X
Ins 432 A (ARH1)	5	4	Sardinian/Sardinian Sardinian	Fs, C143 + 27X
T ₁₅₁₂ G	1	IVS4	Mexican	A protein lacking 26 amino acids from PTB domain
Ins C ₅₉₉	1	6	Japanese	L203 + 16X
	1		English	
Del exon ~1-7		~1-7	English	Null
G ₁₇₇₃ A	1	IVS7	Italian	W249 + 63X

The mutations in *ARH* can be classified into three classes.

Large deletions and insertions: Two *ARH* mutations are reported in this group. In both cases transcription is impaired by large gene deletions or insertions, one of them being an insertion in intron 1 derived from an Alu element (Long Terminal Repeat) of the human endogenous retrovirus HERV-K18. The mechanism by which this insert interferes with *ARH* expression is not known (Wilund et al. 2002).

¹ Nucleotide numbering based on ATG=1,2,3

² G24 + 31X indicates 31 additional amino acids are predicted to be encoded after G24 before a stop codon

Splice site mutations: Two have been reported, with the first affecting a splice-site in intron 1. The predicted protein is not known because of unavailability of mRNA from the patients for analysis (Al-Kateb et al. 2003). The second mutation in this group was reported recently in two siblings with ARH in Mexico. After analyzing the exon-intron boundaries by PCR of genomic DNA, a homozygous sequence change (T→G) was found at the +2 position of the donor splice site of intron 4. The analysis of *ARH* mRNA from the patients revealed two cryptic splice sites, which generated three different transcripts, these being: 1) A minor transcript derived from the use of the normal donor splice site, 2) a second minor transcript which carried a 34-nucleotide deletion in exon 4 and predicted to give a truncated protein lacking 149 amino acids and 3) a major transcript carrying an in-frame deletion of 78 nucleotides in the mature mRNA resulting in a protein that lacks 26 amino acids of the PTB domain. A Western blot analysis of lymphocyte protein extracts from these patients showed the presence of ARH protein of slightly lower molecular weight than normal, which confirmed a translation of the protein lacking 26 amino acids from the PTB domain. This ARH might interact with other receptors besides LDLr showing the functional significance of the mutation (Canizales-Quinteros et al. 2005).

Truncation mutations: Of the 9 mutations reported in this group, five of them lead to a protein lacking the entire PTB domain; these mutations are all found in exon 1. The remaining 4 occur in exons 4, 6 and 7, and comprise nonsense mutations and frame shift insertion or deletion events. The possibility of hot spot mutations in *ARH* is evident because two of the mutations from this group were found in run of nucleotides, these

being a run of Gs in exon 1 and run of Cs in exon 6 (Garcia et al. 2001; Eden et al. 2002; Wilund et al. 2002; Harada-Shiba et al. 2003).

1.4.6 *ARH* Mutation Demographics

Mutations in exon1 and exon 4 are prominent in Sardinian families. These are ARH1 (an 'A' insertion in exon 4) and ARH2 (G→A change in exon1). These mutations introduce stop codons that causes premature termination, immediately before the PTB domain. The high frequency of these ARH gene mutations in Sardinia is likely to be ascribed to a combination of genetic drift and geographic isolation. It also strongly suggests that these mutations were introduced at an early stage in the history of colonization of the island. Previous estimations suggest that approximately 10% of marriages within Sardinians between 1910 and 1969 were consanguineous. The existence of ARH3 (the combination of the two mutations) probably arose from a recombination event between exon 1 and exon 4 in a Sardinian ARH subject who was a compound heterozygote for mutations ARH1 and ARH2 (Arca et al. 2002).

1.4.7 *ARH* mRNA Stability

Most of the fibroblast lines from ARH patients, studied to date, have shown an absence of *ARH* mRNA. This is not unexpected as mRNA transcripts containing premature termination codons are specifically recognized by a special surveillance mechanism and degraded more rapidly than the normal mRNA. This process, called nonsense-mediated mRNA decay, permits cells to eliminate the mutant mRNAs from their translatable mRNA pools, thereby preventing the formation of truncated polypeptides that may overwhelm or disrupt normal cellular mechanisms (Wilund et al. 2002).

1.4.8 *ARH* Polymorphisms

Several silent and non-silent polymorphisms have been reported in the coding sequence of *ARH* (Harada-Shiba et al. 2003; Al-Kateb et al. 2003). While none of these polymorphisms have emerged as important genetic determinants of plasma cholesterol levels in normolipidemic subjects (Hubacek & Hyatt 2004), it is unknown whether any of them have an effect on *ARH* expression or protein stability.

1.4.9 Treatment

Patients with ARH respond to lipid-lowering medications resulting in noticeably greater reductions in plasma cholesterol compared to patients with homozygous FH. Statins produce striking reductions in plasma cholesterol in some ARH patients. The increased LDLr expression induced by the drugs may somehow compensate for the defect in LDL uptake and degradation (Arca et al. 2002).

Different studies have been done on ARH but the cellular function of ARH, the mechanism of the interaction and also the process of LDL uptake is still not well understood. A better understanding of the molecular and genetic mechanisms of this genetic defect may contribute to the management of the disorder.

1.5 PROJECT AIM

The aims of this project were two fold. The first being to confirm the provisional diagnosis of ARH in two patients who presented with the clinical and laboratory features of ARH deficiency. This aim focused on determining the cellular activity of ARH in skin fibroblasts and transformed lymphocytes from the patients and normal controls. Further confirmation of an *ARH* defect was to be sought by sequence analysis of cDNA and

exonic regions of the gene. Any mutation found was to be fully worked up in terms of PCR detection, haplotype determination and family screening. The carrier frequency in the general population was also to be determined by the screening of cord-blood spots from black newborns, which have been anonymously banked at Red Cross Children's Hospital.

University of Cape Town

CHAPTER 2:

MATERIALS AND METHODS

2.1 PATIENTS	28
2.2 CELLULAR INVESTIGATIONS	31
2.2.1 Tissue Culture	31
2.2.1.1 Skin Fibroblasts	32
2.2.1.2 Transformed Lymphocytes	32
2.2.2 LDL-receptor Assay	34
2.2.2.1 LDL preparation	34
2.2.2.2 Protein determination	36
2.2.2.3 LDL binding and internalisation assays	36
2.3 MOLECULAR STUDIES	38
2.3.1 Total RNA Extraction	38
2.3.2 Polymerase Chain Reaction (PCR)	39
2.3.3 Confirmation of the identity of the RT-PCR product	43
2.3.4 Sequencing	44
2.3.5 Mutational Analysis	44
2.3.6 Single Nucleotide Polymorphism (SNP)	46
2.3.6.1 Exon 6 SNP (C→T)	46
2.3.6.2 Exon 7 SNP (G→A)	47
2.3.6.3 SNP frequency analysis	48
2.3.6.4 Haplotype Analysis	48
2.3.7 Western Blotting	49
2.3.7.1 Protein extraction	49
2.3.7.2 SDS PAGE	49

2.1 PATIENTS

All the patients and family members in this study gave informed consent for genetic and molecular studies performed.

PATIENT I

The first patient examined in this study is a Zulu male who was first seen at the Lipid Clinic of the Johannesburg General Hospital in 1993, at the age of 27 years. He had a normal history but was aware of lumps (xanthomata) on the elbows, ankles and hands from the age of 8 years. As an adult he presented to the lipid clinic with;

- Bilateral arcus comealis
- Thickened achilles tendons (Figure 2.1.A) and ligamentum patellae xanthomata
- Tuberos xanthoma on left elbow (Figure 2.1.B)
- Planar xanthoma on right antecubital fossa

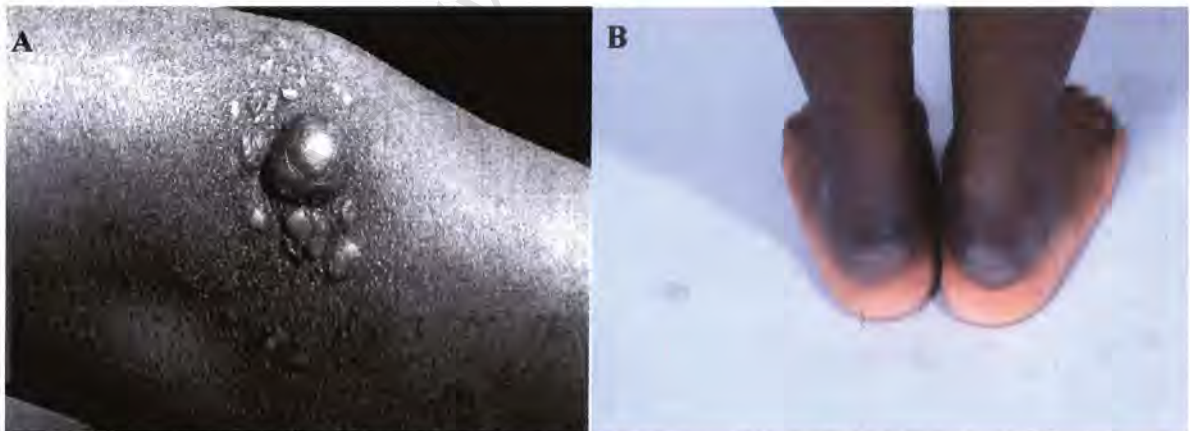


Fig. 2.1: Physical signs of the subject. A) Tuberos xanthoma on elbow B) Thickened achilles tendons.

Plasma lipid levels on presentation are not available but reference to laboratory evaluations in 1992 show plasma total cholesterol of 14 mmol/l with LDLc of 12 mmol/l, fasting triglycerides and HDL-cholesterol were in the normal range. The patient's highest

recorded untreated total cholesterol was 19.4 mmol/l (1996). The pedigree of the family is shown in figure 2.2. There is no family history of xanthomata or any premature heart disease.

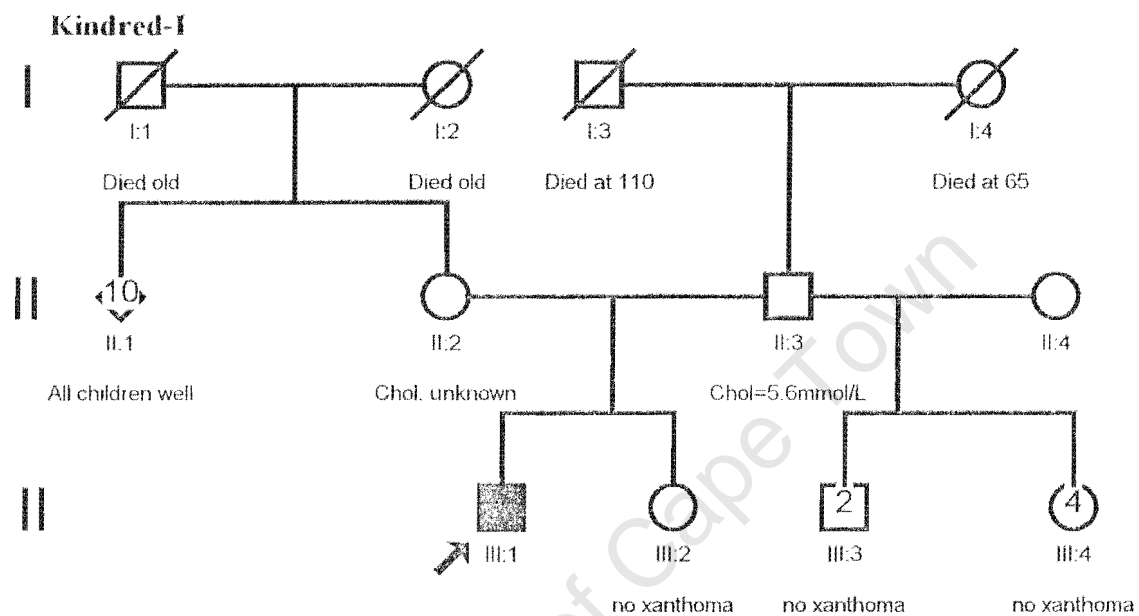


Fig. 2.2: The family tree of Patient I (III.1). Squares, males; circles, females; filled squares, affected individuals; Slash on the signs, deceased.

A diagnosis of homozygous familial hypercholesterolemia was initially entertained for patient I, as this is the most common explanation of an isolated LDLc elevation in subjects of all racial groupings. Of concern however, was the absence of lipid data from his mother, the normal plasma cholesterol level in his father and the absence of early deaths in the family as shown in figure 2.2. He also responded well to statin therapy and typically presented with a 50% reduction in total plasma cholesterol. This is a much better than the average reduction seen in homozygous FH due to LDLr mutations. In the absence of functional assays of LDLr activity, the possibility of FH was explored by sequencing the promoter region and the entire coding sequence of the LDLr gene. This

was carried out by Dr JC Defesche (Dept. of Vascular Medicine Academic Medical Center Amsterdam) in 2003. While no sequence alteration was found in the coding regions, a guanine to thymine substitution (G→T) was found at position 175 of the promoter region. The patient's father was found to be homozygous for the variation and as he presented with a normal lipid profile, the base change was considered to be a polymorphism.

Patient I was recalled at this stage and a skin biopsy obtained for functional assessment of LDLr activity. Blood was also taken for DNA and mRNA isolation, the latter for the purposes of RT-PCR studies.

PATIENT II

The second family is of mixed ancestry and whose pedigree is shown in figure 2.3. The proband in this family is a 5-year-old boy who presented with xanthomata at an early age. He had the clinical signs of homozygous FH with fasting triglycerides of 1.2 mmol/l, total cholesterol of 17.0 mmol/l, LDLc of 15.2 mmol/l, and HDLc of 1.2 mmol/l. His father and mother had slightly elevated LDLc levels of 5.2 and 4.2 mmol/l. There was no other family member with xanthomata and no family member with premature heart disease.

Table 2.1: Plasma lipid levels (mmol/l) of patient II and his parents.

Subject	Total cholesterol	HDLc	LDLc	TG
Proband B	17.0	1.2	15.2	0.9
Father	6.2	0.8	5.2*	0.9
Mother	5.8	1.4	4.2*	0.5

(* Lipid levels above normal but very low for a heterozygous FH)

ARH deficiency was considered as the provisional diagnosis in this kindred and not FH due to LDLr mutation mainly because of the lower than expected LDLc levels of both parents and the absence of premature coronary artery disease in family members. A skin biopsy was taken from patient II and his father for cellular studies while blood was collected for molecular analysis.

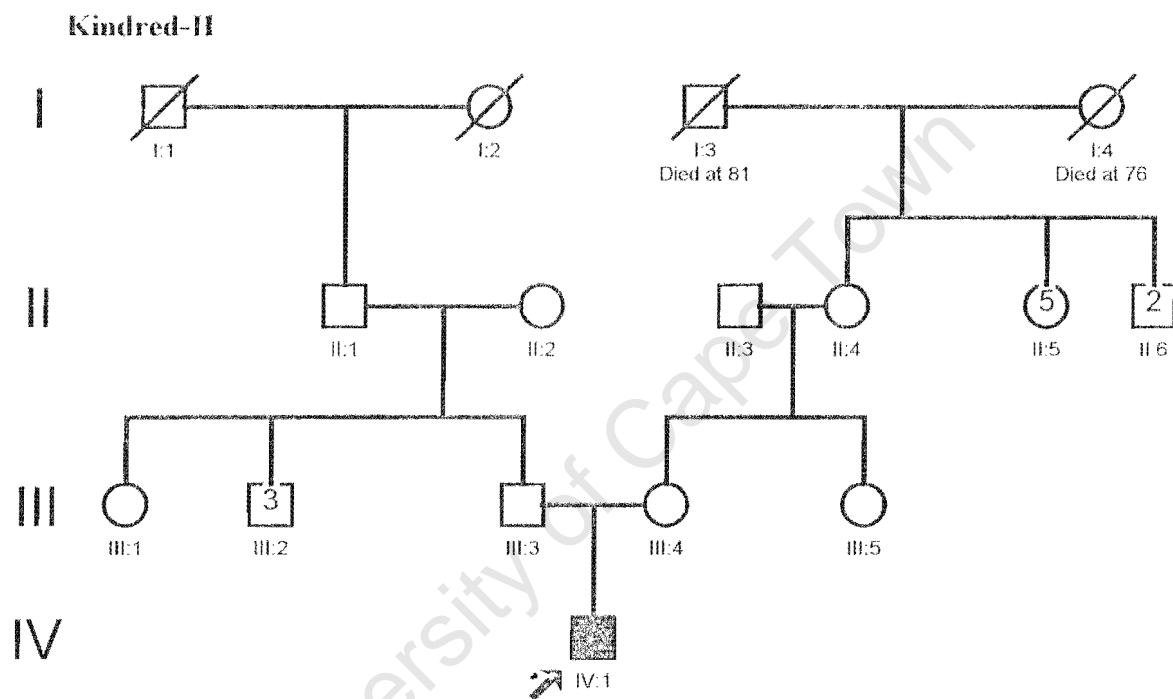


Fig. 2.3: The family tree of patient II (IV.1)

2.2 CELLULAR INVESTIGATIONS

2.2.1 Tissue Culture

The medium used for all cell cultures contained Dulbecco's Modification of Eagle's Minimum Essential Medium (DMEM, Gibco); with 10% heat inactivated fetal calf serum (HIFCS), penicillin G (60µg/ml) and streptomycin sulphate (100µg/ml). The medium was

stored at 4°C, and was warmed at 37°C before use and was frequently checked for contamination by swirling the medium.

2.2.1.1 Skin Fibroblasts

Fibroblasts were either obtained from frozen down cultures or grown from forearm skin biopsies. Monolayer cultures were grown in 75 cm² flasks and kept in a humidified incubator (95% relative humidity, 5% CO₂). The cells were checked for a growth every other day using an inverted microscope and phase-contrast optics. When the cells were confluent they were washed twice with phosphate buffered saline (PBS), and either harvested by mechanical scraping or dissociated with 5ml 0.05% trypsin (Difco)/0.02 EDTA (Merck) solution for 5 min at 37°C. In either case, the pellet was re-suspended in fresh medium for re-seeding of flasks, with the new passage number being recorded on each flask. Cells were also frozen down for later use by re-suspending the trypsinized cells (~1 × 10⁶ cells/ml) in a medium containing dimethylsulphoxide (10%; DMSO) and transferred into small vials. The vials containing cell aliquots (1ml each) were placed upright in a -80°C freezer and later transferred to liquid nitrogen for long term storage. For re-use, the vials were thawed quickly at 37°C and suspended in 10ml full medium to neutralize the DMSO. The mix was then spun at 2000rpm for 2 min and cell pellet re-suspended in 10 ml of full medium in a 75 cm² tissue culture (TC) flask (Paul 1975; Sly & Grubb 1979).

2.2.1.2 Transformed Lymphocytes

Lymphocytes were transformed with Epstein-Barr virus (EBV). Transformation medium, containing EBV, was prepared from B95/8 cell cultures. The cells were seeded (1 × 10⁶ cells/ml) into 75 cm² TC flask and left to grow until the medium was well conditioned for

about three days. This medium was collected by pelleting the B95/8 cells and filtering the supernatant through a 0.22 μ m filter (John et al. 1991).

Heparinized blood samples (10ml) were collected from patient II, his father and two normal controls. Buffy coats were aspirated in a volume of 200 μ l after spinning the blood samples for 10min at 3000rpm. These were layered on 5ml of lympho prepTM Ficoll hypaque (Sigma), and centrifuged for 20 minutes at 3000rpm. The clear white cell band was removed and washed with culture medium twice at 1000rpm, 10 minutes in the first instance and 5 minutes in the second. The pelleted cells were re-suspended in a mix of transformation medium (5ml) and phytohaemagglutinin (PHA 1%), transferred to a 25cm² TC flask and placed in the CO₂ incubator. Small volumes of medium (free of penicillin & streptomycin) were added to the cells until they were ready to passage into 75 cm² flasks. Cell growth was assessed by visual inspection and maintained by regularly discarding about 50% of the cells and medium, and replenishment by adding the same amount of fresh medium. The transformed lymphocytes normally grow as suspensions of large clumps that are easily broken up by aspirating gently several times. For cell-bank purposes, cells in a 75cm² flask were washed twice using full medium and then re-suspended in a freezing medium (DMEM with 20% HIFCS & 10% glycerol). Aliquots were quickly frozen using a Cryoson freezing-down machine that is connected to a liquid nitrogen source. Upon the completion of the freezing cycle the vials were stored in appropriate boxes in a liquid nitrogen storage tank, with one vial being thawed to check on viability during the freezing process (John et al. 1991; van der Westhuyzen et al. 1984). Before freezing, the cell suspension was checked for Mycoplasma contamination by putting a drop of the suspension on a slide and fixing it with methanol. A working

fluorescent stain (Appendix A) was applied to the slide for 30sec, with the slide being subsequently washed under tap water. The cells were then observed on a Zeiss 510 Meta microscopy for mycoplasma staining (Chen 1977).

2.2.2 LDL-receptor assay

2.2.2.1. LDL preparation

Plasma was collected (125ml), centrifuged (6300g) at 10°C for 11 minutes and aliquoted into two tubes. The supernatant was decanted and the total volume was measured. Protease inhibitors aprotinin (2ml in 500ml of plasma from a 250×stock concentration [10000U/ml]) and phenylmethylsulfonyl fluoride (PMSF; 1ml in 500ml of plasma from a 500× stock concentration [100mM]) were added to prevent protein degradation. For the purposes of preparing labeled LDL, the fluorescent probe Dil (1,1'-dioctadecyl-3,3,3'-tetramethylindocarbocyanine perchlorate; 6mg/ml of plasma) was added to one of the tubes, which was then incubated overnight at 37°C (Pitas et al. 1981). KBr (49.5g/100ml) was added to both tubes to raise the density to 1.3g/ml; dissolution of the KBr was achieved by swirling.

Quickseal tubes were filled with 27ml of Saline-EDTA (NaCl 0.15 M & 0.01% EDTA) and 13ml of $d=1.3\text{g/ml}$ plasma was carefully layered at the bottom of the tubes without mixing the saline and plasma. The tubes were properly sealed and centrifuged using the following parameters: minimum temperature of 10°C, maximum of 20°C, speed of 49,500rpm, time 2,5hours, slow acceleration and deceleration. After centrifugation the tubes were removed and the different bands of lipoproteins were marked as VLDL, LDL, HDL & Lipoprotein Deficient Plasma [LPDP] (Figure 2.4).

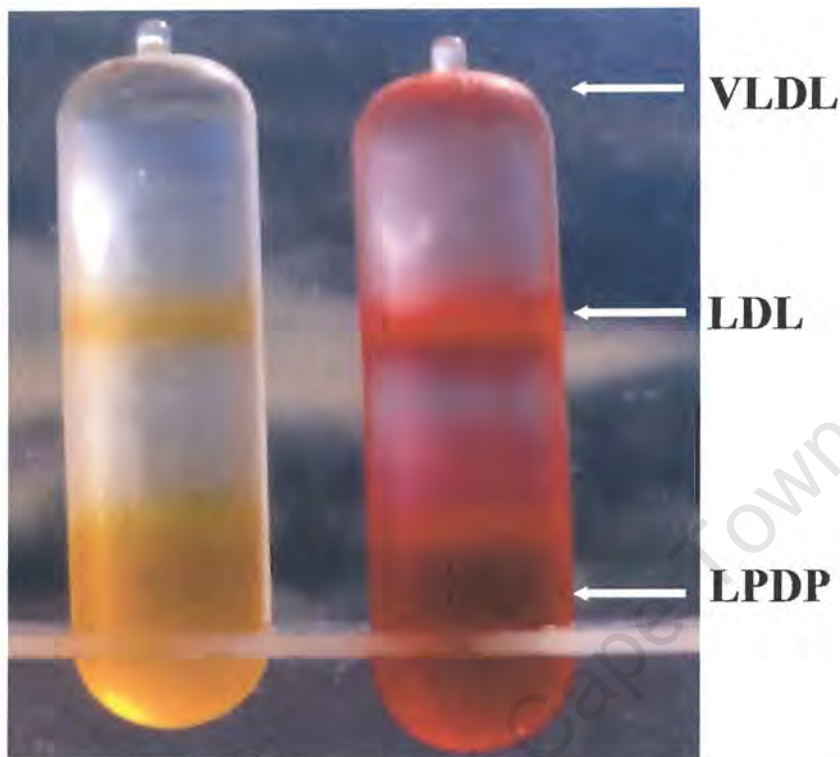


Fig. 2.4: Separation of plasma lipoproteins by density gradient ultra centrifugation. The left tube contains the unlabelled LDL while the right one contains the DiI fluorescent probe.

To slice out the marked LDL band, the tubes were put into the slot of the cutting device and positioned such that the top of the LDL band was just above the surface of the cutting blade. After cutting, the LDL layer was collected by aspiration using a pasteur pipette. The very bottom layer, which represented the lipoprotein deficient plasma, (LPDP) was also collected by aspiration. KBr was removed by dialysis using dialysis tubing (exclusion size 10 000da; 6mm and 43mm diameter) and cold saline/EDTA (200ml for the LDL prep and 400ml LPDP prep), with stirring and a change of dialysis fluid every 4hrs for 16 hrs (Chung et al. 1980).

After dialysis the LPDP was heat inactivated at 56°C for 30 minutes and sterilized through 0.45µm filter for further use (Chung et al. 1980). The protein concentration of the LDL was determined as described below and stored at 4°C until further use.

2.2.2.2 Protein Determination

Protein concentrations were determined by using Markwell's modification of the Lowry method as this gives optimal detection of proteins from a lipid environment (lipoproteins and membranes) (Markwell et al. 1978). Bovine serum albumin (BSA 1mg/ml) was used as a standard and was prepared by dilution from 0 to 250µg of protein in 250µl of distilled water in calibration tubes. Both LDL preparations were assayed in volumes of 250µl at 1:10 and 1:25 dilutions in distilled water. Solution C (750µl: A mix of solution A and B, Appendix A) was added to the calibration tubes, which were incubated at room temperature for 15 minutes. After the incubation, 75µl of diluted Folin Ciocalteu solution was added to the tubes, forcibly and inverted to give rapid mixing. The absorbance was read at 750nm after an incubation of 45 minutes at room temperature for all specimens. A standard curve was plotted and the unknown protein concentration was calculated.

2.2.2.3 LDL binding and internalisation assays

Sub-confluent fibroblast cells from the patient-I and two normal controls were trypsinized and re-suspended in 4ml of full DMEM medium. The cells were then set up in 6×60mm dishes, in duplicate in two parallel arms, at approximately 100,000 cells per dish and were incubated in 5% CO₂ in a humidified incubator for a period of two days. On day three the medium was changed and at ¾ confluency, LDLr gene expression was up-regulated by placing the cells in 2ml of DMEM medium (no FCS; 60 :g/ml penicillin G, 100 :g/ml streptomycin), containing LPDS (5mg/ml). Cells in each arm of each

culture were exposed to DiI-labeled LDL at 10 μ g/ml, with cells in arm B additionally getting unlabelled LDL in 20 fold excess, to determine non-specific binding and internalisation. After 5hrs of incubation the cells were kept on ice for 10min and then washed 7 times, 4 \times with albumin-PBS (1g albumin per 500ml PBS for a final concentration of 0.2% albumin) and 3 \times with PBS. Lysis buffer (0.1%SDS/1M NaOH) was added to each dish and left overnight at 4 $^{\circ}$ C and this reagent allowed the direct determination of DiI fluorescence intensity (FI) and cell protein concentration in the same sample of lysed cells. FI was measured in the lysate on an opaque white plate using Carry Eclipse program with the following parameters; Florescence excitation wavelength (Ex): 520nm, emission wavelength (Em): 580nm, Ex slit: 5nm, Em slit: 10nm and PMT detector voltage: medium (Carry Eclipse software manual www.varian.com). The florescence for both arms (labeled and unlabelled) was read and the difference was taken to rule out non-specific binding. The same samples were used to measure protein concentration using Markwell's modification of the Lowry method. The DiI-LDL florescence data for each sample was then expressed per mg of its respective cell protein concentration (Figure 2.5) (Pitas et al. 1981).

This is an example of the calculation

Arm 1 (Hot)

Protein concentration=0.243mg/ml FI=250units/ml (total volume per dish is 1 ml)

FI per mg cell protein= 250units/0.243mg per dish

$$=1028 \text{ units per mg}$$

Arm 2 (Cold)

Protein concentration=0.23mg/ml FI=18.4units/ml (total volume per dish is 1 ml)

FI per mg cell protein= 18.4units/0.23mg per dish

= 80 units/mg

Final LDLr specific FI per mg cell protein = Arm 1- Arm 2

= 1028-80

= **948 units/mg**

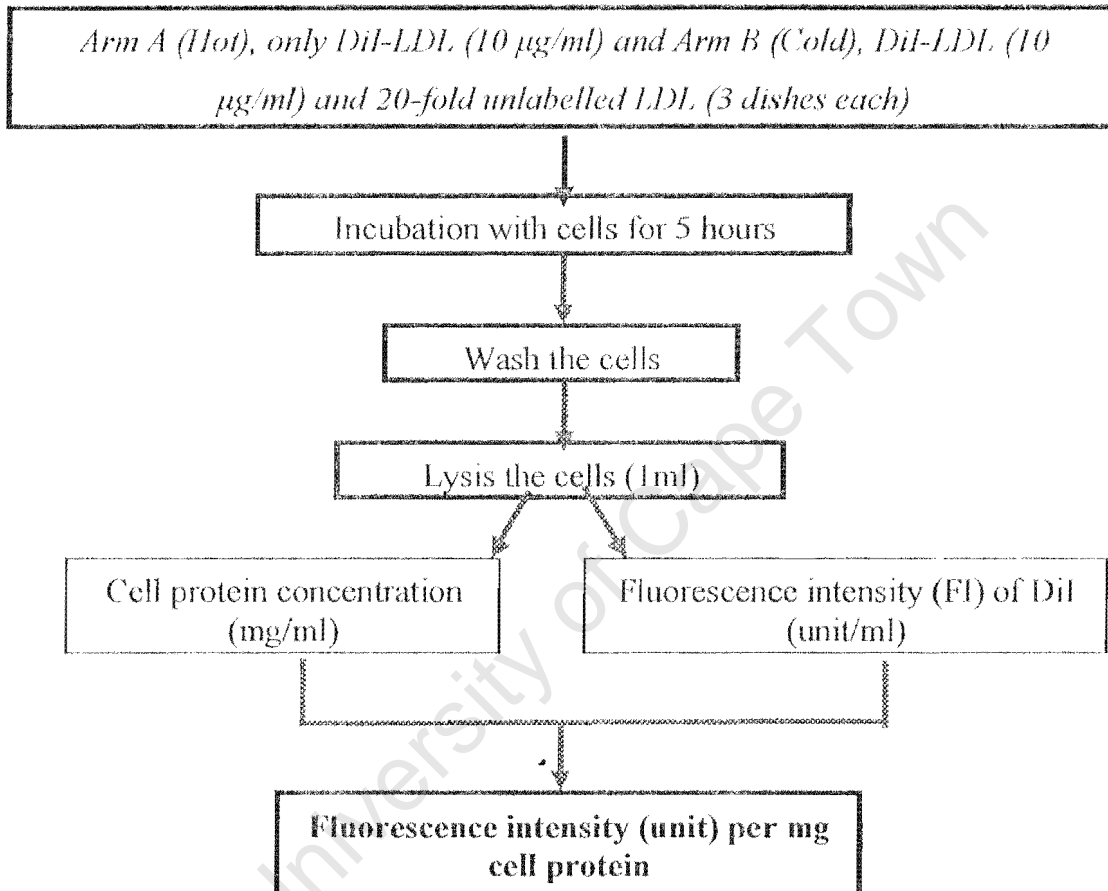


Fig. 2.5: A flow diagram of steps in the spectrofluorometric assay for Dil incorporated LDL uptake.

2.3 MOLECULAR STUDIES

2.3.1 Total RNA Extraction

Throughout the RNA extraction, a ribonuclease (Rnase) free environment was maintained by using double distilled deionized water that was pretreated with diethylpyrocarbonate (DEPC-ddH₂O) and by wiping all the surfaces with Rnase-Away™

(Gibco-BRL). Total RNA was extracted from skin fibroblast cells using the RNA extraction reagent TRIpure™ (Roche). The cells were lysed in the TRIpure reagent by pipetting a few times with a pasteur pipette and incubating at room temperature for 5 minutes to ensure complete disassociation of nucleoprotein. Chloroform (0.2ml per 1ml of TRIpure™ used) was added and the sample was shaken vigorously for 15 seconds. After incubation (15 min at room temperature) the sample was centrifuged at 1200g for 15 minutes at 4°C. The colorless upper phase that contains the total RNA was transferred to a clean Eppendorf tube and precipitated by the addition of isopropanol (0.5ml per 1ml of TRIpure™ used). The tubes were incubated at room temperature for 10 minutes and centrifuged at 1200g (10 minutes at 4°C). After discarding the supernatant, the pellet was washed with 1ml of 75% ethanol in DEPC-ddH₂O and centrifuged at 7500g for 5min at 4°C. The RNA pellet was air dried for 10 minutes and then resuspended in approximately 40µl DEPC- ddH₂O. The sample was then aliquoted to avoid repeated thawing and freezing and stored at -70°C. The concentration and purity of the RNA were determined by measuring the absorbance at 260nm and 280nm (OD₂₆₀ and OD₂₈₀) and by UV-light viewing after agarose gel electrophoresis and ethidium bromide staining (Chomczynski & Sacchi 1987).

2.3.2 Polymerase Chain Reaction (PCR)

All the primer pairs used in this study were designed using the Primer 3 web tool program except the primers for exon 1. The sequence of this exon does not appear in any of the nucleotide databases and it was necessary to use the primer sequences reported in a prior publication (Arca et al. 2002). The Primer 3 programme can be accessed at the following Web address:

(<http://www-genome.wi.mit.edu/cgi-bin/primer/primer3 WWW.cgi>)

The specificity of the designed primers was checked by a BLAST search of GenBank using a program called (NCBI BLAST TOOL) (<http://www.ncbi.nlm.nih.gov/BLAST/>).

Mismatch primers for the introduction of restriction sites in mutation and SNP analysis, were designed manually.

Table 2.2: Primers used for amplifying the coding sequence of *ARH*.

<i>ARH</i>	Sequence 5' > 3'	PCR product (bp)
cDNA F	TCCTGACGGAGTTTTGGCTG	1027
cDNA R	TTTGGTCCATCACGTGTCAGG	
Exon 1 F	TGTCCTCTCTCGCCCCGCCCGT	~380
Exon 1 R	GCGCCGGTGAGGCCCGCCTCT	
Exon 1 F (Mismatched)	CCAGCTTGGCCAAGCAGAGCCGG	
Exon 6 F	ACATGTTGTGCCTTGGTCCTGC	120
Exon 6 R (Mismatched)	GTCAGGACTCACAGCCTTTCAAG	
Exon 7 F	GAAGAGGCTGATCTCCCACTGA	285
Exon 7 R	CCAGAGGCCACAAGTCACG	

Reverse transcriptase PCR (RT-PCR)

The cDNA template for RT-PCR is synthesized from total RNA by reverse transcription.

The reaction components and the respective concentrations are given below in the Table 2.3 and 2.4.

1) RNA target and Primer mix

For each 20µl of reverse transcription (RT) reaction, a mix of RNA template and primers in a sterile nuclease-free tube was prepared. Amplification of cDNA using only oligoδT primer was not successful. This was repeated by using random hexamer primers in conjunction with the oligo δT primer. The components and the concentrations of the

RNA template and primers mix taken are shown in table 2.3. The RNA mix was then incubated at 70°C for 5 minutes and quickly chilled at 4°C for 5 minutes.

Table 2.3: Reaction mix for the RT-PCR

Component	Stock concentration	Final concentration
Primer oligo δT (Promega)	500 pmol/ μ l	20 pmol/ μ l
Primer Hexamers (Promega)	500 pmol/ μ l	20 pmol/ μ l
RNA	0.232 μ g/ μ l	1 μ g

2) Reverse Transcription Mix

For each 20 μ l RT reaction, the components mentioned in table 2.4 were combined.

Table 2.4: Reverse Transcription mix component concentrations.

Component (Promega)	Stock concentration	Final concentration
ImProm-II TM Reaction Buffer	5 \times	1 \times
ABO196 δ NTP Mix	5 mM	0.5 mM
MgCl ₂	25 mM	3 mM
Reaction buffer	10 \times	1 \times
rRNasin® Ribonuclease Inhibitor	40 u/ μ l	20u
ImProm-II TM Reverse Transcriptase	-	1 μ l/reaction

After vortexing the reverse transcription mix, the RNA mix was added and reverse transcribed using a Gene Amp PCR system 9700 thermocycler (Perkin-Elmer). The reaction conditions are mentioned in table 2.5.

Table 2.5: The RT-PCR reaction conditions.

Cycles	Temperature (°C)	Time (Min)
Annealing	25	5
Extension	42	60
Heat inactivation of RT	70	15

The cDNA was then amplified using specific primers and the PCR components are mentioned in table 2.6.

cDNA primers: Forward: 5' -TCCTGACGGAGTTTTGGCTG -3'
 Reverse: 5' -TTTGGTCCATCACGTGTCAGG -3'

Length of PCR product: 1027 bp.

Table 2.6: PCR component concentrations.

Component	Stock concentration	Final concentration or volume
Q solution	5×	1×
Primers each	20 pmol/μl	0.4 pmol/μl
δNTP Mix	2.5 mM	0.2 mM
Reaction buffer (15 mM MgCl ₂)	10×	1×
HotStar Taq polymerase	5u/μl	0.025 u/μl
cDNA	-	1μl

Optimization of cDNA amplification

Exon 1 of ARH is very GC rich and proved to be difficult to amplify under normal PCR conditions. The following optimizing alterations were introduced to enable successful amplification.

I. HotStarTaq™ DNA Polymerase: This Qiagen product has been developed to provide for “hot start” PCR, which provides for greater specificity of amplification in general, but is particularly suited to amplification of difficult templates, such as GC rich regions (HotStarTaq PCR handbook). The components that provide for the specificity are:

- **Q-Solution:** This is an innovative PCR additive that facilitates amplification of difficult templates by modifying the melting behavior of DNA.

- **HotStarTaq Polymerase:** This enzyme is provided in an inactive state with no polymerase activity at ambient temperatures. It is activated by a 15-minute, 95°C incubation step, which can easily be incorporated into existing thermal cycling programs.

II. Touchdown PCR

Touchdown PCR involves decreasing the annealing temperature (0.5°C every second cycle in this case) to a ‘touchdown’ annealing temperature, which is then used for the remaining cycles. This allows for higher specificity of amplification during the initial cycles and preferential amplification of the early templates at the lower Tms.

Table 2.7: Touchdown PCR reaction conditions

Cycles	Denaturation Temp (°C) /time	Annealing (°C)/time	Temp	Extension Temp (°C)/ time
First hold	95/60 sec	-		-
First 10 cycles	94/30 sec	71-66/ 45 sec (decreases 0.5 °C per cycle)		72/ 1min
Second 30 cycles	92/30 sec	66/45 sec		72/1 min
Final hold	-	-		72/7 min

2.3.3 Confirmation of the identity of the RT-PCR product

The amplified cDNA was confirmed by digestion with *Taq I* and comparing the fragment patterns on agarose gels with the pattern predicted from the *ARH* cDNA sequence. This was achieved as follows: PCR product (10µl) was digested at 65°C for 5hrs with 5 units of *Taq I* (5U/µl), 2µl of 10× restriction buffer and 7µl water (Total volume 20µl). *Taq I* has a **T/CGA** recognition site and cuts the PCR product three times giving four bands with lengths of 91, 108, 356 and 472 bp. After the incubation, 4µl of gel tracking dye (Appendix A) was added to stop the reaction and to increase the density of the PCR

product. The bromophenol blue in the dye acts as a tracking dye to follow the progress of the electrophoresis. The mix was then loaded onto a 4% agarose gel that was melted in tris-acetate/EDTA (TAE) buffer (Appendix A) with 2µl of ethidium bromide (10mg/dl). To check the digestion, electrophoresis was performed at 100 volts/20cm for 30 min. and the gel was visualized using ultraviolet light (Sambrook et al. 1989).

2.3.4 Sequencing

PCR products from the subject and control were sent for cycle sequencing to Dr. Patricia Owen (University of Cape Town, department of Chemical Pathology) who used the ABI BigDye sequencing kit and ABI Prism sequencing apparatus.

2.3.5 Mutational Analysis

The *ARH* cDNA product was sequenced using forward and reverse primers and about 650bp was clearly read from both ends. A comparison of the cDNA sequence from the patient with that of the *ARH* sequence lodged in GenBank revealed the patient to be homozygous for a 'G' insertion in a run of 7 'G's in exon 1. This insertion mutation was authenticated by repeat sequencing of PCR product from a different amplification run in addition to sequencing PCR amplified exon 1 from both parents of the patient.

Exon 1 primers: Forward: 5' -TGTCCTCTCTCGCCCCGCCCCGT- 3'
Reverse: 5' -GCGCCGGTGAGGCCCCGCCTCT- 3'

Length of PCR product: ~380 bp.

A mismatch forward primer was designed to confirm the homozygous 'G' insertion found in exon 1 from the subject and for further frequency analysis with restriction enzyme. This was done by a T→C substitution immediately preceding the run of 'G's in

the amplified product. This created a *Bgl I* cutting site (GCCN₃GGC) in the PCR product of normal *ARH*, and which would be destroyed by the ‘G’ insertion mutation.

Amplification was effected by using the mismatch forward primer with the previous reverse primer, giving a PCR product of ~170bp.

5'...GCTGGGGGGGGC...3' W 5'...GCGGGGGGGGC...3'

Previous sequence

Mismatched sequence

Exon 1 mismatched forward primer:

5' -CCAGCTTGGCCAAGCAGAGCCGG -3'

Table 2.8: Digestion mix for exon 1 mutation detection.

Components	Stock concentration	Final concentration or volume
PCR product		10µl
SuRE/Cut Buffer H	10×	1×
Distilled water		7.5µl
Restriction enzyme (<i>Bgl I</i>)	10 U/µl	5U

Complete digestion of the PCR product of the normal allele was found to occur on overnight digestion at 37°C. This gives a two-banded pattern on electrophoresis with fragment lengths of 150 and 20 bp.

The estimated frequency of this mutation in the black community was sought by analyzing blood spots from 400 newborns collected as part of a newborn bloodspot banking exercise carried out at the Red Cross Children's Hospital, Cape Town. Small disks (3mm) were punched from the blood spots and were fixed with methanol for 5 minutes. After air-drying, the blood spot was incubated in 60µl of distilled water (for 15 minutes at 95°C), with 5µl of the mix being taken for PCR. The same PCR components' concentrations (Table 2.6) and conditions (Table 2.7) were used except the final number of cycles was increased by 10 as the blood spots contain low number of templates.

The same PCR conditions as table 2.7 were used except the touchdown annealing temperature was lowered from 71°C -66°C to 68°C-63°C. The PCR mix was prepared in a final volume of 25µl and was amplified with a Gene Amp PCR system 9700 thermocycler (Perkin-Elmer). PCR product (12µl) was digested at 55°C for 3hrs with *Bsi* *YI* after adding five units of enzyme (0.5µl of *Bsi* *YI*), 2µl of 10× restriction buffer M and 5.5µl distilled water (Total volume 20µl). The 'C' to 'T' variation changes the first base pair of *Bsi* *YI*'s recognition site and abolishes the digestion while the normal sequence gives two bands with 102 and 18 bp after complete digestion.

2.3.6.2 Exon 7 SNP (G→A) NCBI ss# 35032448

The third base change found in patient I occurred in exon 7 at position 654 of the cDNA (ACG→ACA). This SNP is silent, as it does not alter the amino acid sequence, and is also novel, as it does not appear in any SNP database. The G→A variation in this exon was easily detected by restriction digestion using *Alu* *I*. This enzyme has a recognition site of AG/CT and cuts the fragment that contains the 'A' allele giving two fragments, of 201 and 84 bp (O/N (@ 37°C) (Table 2.8). *

The following primers were designed for the purposes of amplifying exon 7. These primers were used to determine the frequency of the SNP and to allow for haplotype determinations.

Primers: Forward: 5' -GAAGAGGCTGATCTCCCACTGA -3'
Reverse: 5' -CCAGAGGCCCAACAAGTCACG -3'

Length of PCR product: 285 bp.

The PCR component concentrations and the conditions used are shown in table 2.9 and 2.7 respectively.

2.3.6.3 SNP frequency analysis

The frequency of the two polymorphisms in black individuals was determined by amplifying both exons from a 100 newborn bloodspots with subsequent restriction enzyme digestion.

Hardy-Weinberg analysis

The observed and expected genotype frequency of both SNPs were estimated from the allele frequency using Hardy-Weinberg analysis which states that in a randomly mating population the proportions of genotypes aa, ab and bb are given by the binomial expansion of $(p+q)^2$ where p= the frequency of allele a and q= the frequency of allele b (since they add up to 1, $q = 1-p$).

$$\text{Total probability} = p^2 + 2pq + q^2 = 1$$

The agreement between observed and expected results were tested with the χ^2 test

$$\chi^2 = \sum (\text{Observed} - \text{Expected})^2 / \text{Expected}$$

2.3.6.4 Haplotype Analysis

The mutation found in our patient I has been reported previously in a 10 year old Iranian child from a consanguineous union (Garcia et al. 2001). The disparate ethnicities of the two patients would suggest that these two mutational events are mutually exclusive and have occurred at different times in the human lineage. An indication of this would be the finding of differing haplotypes for the two mutant alleles. To explore this we compared haplotypes as derived from the two SNPs reported above. DNA from the Iranian patient and his family was kindly provided by Prof E Leitersdorf, University of Amsterdam, Holland, and was genotyped as described previously. Haplotypes were constructed and compared between the Iranian and the black patient.

2.3.7 Western Blotting

2.3.7.1 Protein Extraction

Fibroblast cells from patient I, a normolipidaemic control and a patient with homozygous FH were seeded at a density of 150 000 cells per 100mm dish and cultured in DMEM (10% FCS). The medium was changed on day 3 and day 5, with the cells being harvested on day 7 by scraping from the dish surface and collected them by centrifugation at 3,000 rpm for 10 minutes.

EBV-transformed lymphoblasts from patient II and his father were seeded at $\sim 1.0 \times 10^6$ cells/ml in DMEM (10% FCS) in 75 cm² flasks. The cells were harvested when they were confluent by centrifuging the medium at 3000 rpm for 10 minutes.

All the cells from the 5 subjects were washed twice with 10ml of PBS (3,000 rpm for 10 minutes) and lysed in 200:1 of pre-chilled lysis buffer containing Tris (50 mM), CaCl₂ (2mM), NaCl (80mM), Triton X-100 (1% [v/v]), pH 8.0 with 1:100 dilution of protease inhibitor mix (Sigma). After homogenizing the mix, they were placed at 4°C for 30 minutes. Lysates were then centrifuged in a microfuge for 15 minutes at 10 000 rpm at 4°C and the supernatant was collected (Arca et al. 2002).

2.3.7.2 SDS PAGE

The Protein concentration of the lysates was determined using the Bio-RAD protein assay reagent and 50:µg of protein was fractionated by electrophoresis (200 V/16cm, 3 h) on a 10% polyacrylamide-SDS gel (Appendix A). The gel was transferred (Appendix A) to a nitrocellulose membrane at constant voltage of 18v/16cm with 100mA current at 12°C overnight (Transphor user manual, Amersham Biosciences). After the transfer was

completed the membrane was stained with Ponceau S stain (500mg/100ml of 1% acetic acid) to check the transfer efficiency.

After putting the membrane in 20ml of blocking buffer (Appendix A) for an hour the membrane was incubated with primary antibody (1:2500) for an hour. This primary antibody was affinity-purified rabbit anti-ARH with a peptide containing residues 180–308 of ARH and was kindly given by Prof. Linton Traub University of Pittsburgh School of Medicine, Pittsburgh (Mishra et al. 2002). The membrane was then washed four times in 20 ml of TBST buffer (Tris Buffered Saline with 0.1% Tween 20) and was incubated in Horseradish peroxidase-linked donkey antibody against rabbit peptides. The enhanced chemiluminescence Western blotting detection kit was used for signal detection.

University of Cape Town

CHAPTER 3:

RESULTS

3.1 PATIENT I	52
3.1.1 LDL receptor function; binding and internalisation assays	53
3.1.2 DNA Studies	53
3.1.2.1 RT-PCR	54
3.1.2.2 Confirmation of the RT-PCR product	55
3.1.2.3 Sequencing	56
3.1.2.4 Mutational Analysis	57
3.1.2.5 Single Nucleotide Polymorphism (SNP)	59
3.1.2.5.1 Exon 6 SNP (C>T)	59
3.1.2.5.2 Exon 7 SNP (G>A)	61
3.1.2.6 Haplotype Analysis	63
3.2 PATIENT II	67
3.2.1 DNA screening	67
3.2.2 SNP analysis	68
3.2.3 Western Blotting	70

3.0 TISSUE CULTURE

Fibroblasts were obtained from frozen down cultures or grown from forearm skin biopsies. No difficulties were experienced with these cultures, which were successfully utilized in LDL uptake assays, Western blotting experiments, and for RNA extraction.

EBV-transformation of lymphocytes was successful for patient II only. Several attempts were made to culture transformed cells from Patient I. In each case the transformation appeared to have been successful, but growth soon ceased with the cells dying within a few weeks. Several troubleshooting approaches were tried, but without resolution. These included the use of well-conditioned medium from confluent healthy lymphocytes, and the increase of FCS to 20%. Contact with other groups working on ARH biology, confirmed this phenomenon of poor viability, but it is clear that not all ARH cell lines are equally affected; there is no published data on this difficulty with transformed lymphoblasts from ARH patients. Patient I blood samples were not transformed in isolation but in parallel with bloods from family members and controls. All these other samples were successfully transformed.

3.1 PATIENT I

Patient I presented with clinical signs of homozygous FH, no family history of premature heart disease, and was provisionally diagnosed with ARH. Cellular and molecular assays were done to confirm the diagnosis.

3.1.1 LDL receptor function; LDL internalisation assays

An assessment of the LDL uptake was only carried out on cultured fibroblasts of patient I due to the failure of EBV transformation lymphocytes to survive; fibroblast lines from two normolipidemic subjects were used as controls.

Uptake studies were carried out on two occasions, with each result being the mean of triplicate dishes. These results are shown in Table 3.1 and are given as fluorescence units per mg cell protein, where they represent both surface bound and internalised LDL.

Table 3.1: Results of the quantitative spectrofluorometric assay for cell-associated DiI-LDL.

Subject	LDLr specific FI per mg cell protein	Average FI per mg cell protein
Patient I	948 811	879.5*
Control 1	788 809	
Control 2	841 1098	884

* Equivalent to 99.4% of control values

The patient's results were averaged and expressed as a % of the fluorescence readings from the control cell lines. This gave patient I an LDL uptake rate which was equivalent to 99.4% of the averaged control cell results.

3.1.2 DNA studies

The search for the mutation underlying ARH in this gene was carried out by sequencing studies following RT-PCR and then by exon amplification of *ARH*.

3.1.2.1 RT-PCR

Amplification of the entire coding sequence of *ARH* presented an early challenge, as the 5' untranslated region of the gene was not available in any sequence database. Also, all the published PCR primer sequences overlapped the ATG start codon; their use would thus have resulted in the loss of some coding sequence. The missing 5' untranslated sequence was located by a BLAST search of the GenBank EST database. Three independent EST sequences were found to contain the required 5' *ARH* data. Sequence was taken from one of these (EST; AL552655.1), which allowed for the design of a forward primer, 50bp 5' of ATG start codon. The chosen primer set thus gave an amplicon of 1027bp in length and which carried both 5' and 3' untranslated cDNA sequence. RT-PCR proved to be difficult under normal PCR conditions as *ARH* is very GC rich (69%), resulting in numerous non-specific bands. In order to remove these non-specific bands and to allow for complete denaturation of the template, a special hot-start *Taq* polymerase enzyme and touch down PCR method was used as discussed in chapter 2. The progression from multiple non-specific bands to predominance of the *ARH* amplicon is shown in figure 3.1.

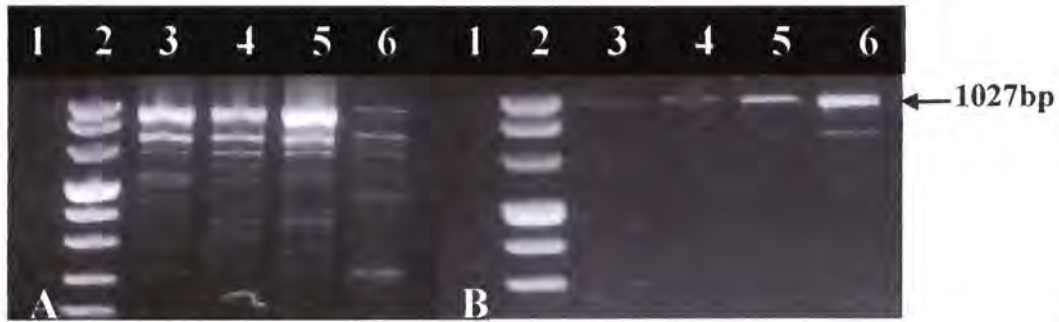


Fig. 3.1: Amplification of *ARH* cDNA following RT-PCR. A) Standard *Taq* polymerase and PCR conditions B) *Hotstar Taq* enzyme with touch down PCR. In both cases Lane 1, blank; Lane 2, DNA molecular weight Marker VIII (Roche); Lane 3-6 PCR product from 32, 34, 36 and 38 final number of cycles respectively.

3.1.2.2 Confirmation of the PCR product

Given the presence of so many non-specific bands it became necessary to confirm the *ARH* identity of the major band, which was ~1000bp in size. This was carried out by comparison of the *Taq I* restriction pattern with that of the *ARH* sequence. *Taq I* cuts the cDNA of *ARH* three times giving four bands with lengths of 91, 108, 356 and 472 bp. After the incubation of the PCR product with the enzyme, the mix was run on a 4% agarose gel to visualize the digestion reaction.

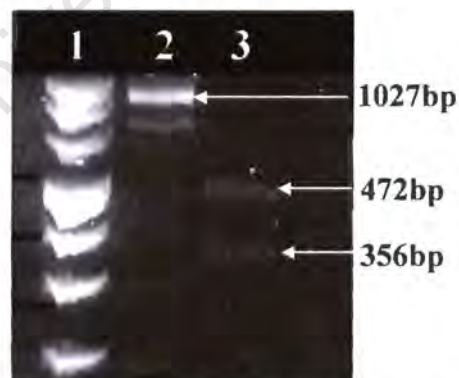


Fig. 3.2: Confirmation of *ARH* sequence in the PCR product by restriction mapping using *Taq I* digestion. Lane 1, Marker VIII (Roche); Lane 2, undigested RT-PCR product (1027 bp); Lane 3, digested RT-PCR product (472 & 356 bp). The 91 and 108bp digestion bands are not shown. The doublet bands are artifactual. Digestion products were run on a 2% agarose gel.

Restriction enzyme digestion of the RT-PCR product confirmed the identity of the band as the three hours of incubation with the enzyme gave the expected cDNA fragment pattern (472 and 356 bp) (figure 3.2). The other two bands were not identified in the gel because of the small size of the fragments (91 & 108 bp).

3.1.2.3 Sequencing

Both strands were sequenced with about 600bp being readable from each side. The sequencing trace of the patient's cDNA was analyzed by blasting it with the normal *ARH* sequence in the NCBI database. The blast revealed three changes. The first one is an insertion of 'G' in a string of seven 'G' residues from nucleotide 65 to 71 of the cDNA (Nucleotide numbering based on ATG=1,2,3). Two single nucleotide polymorphisms (SNPs) were also found, one in exon 6 and one in exon 7 and they are discussed in their relevant sections.

To confirm the 'G' insertion in exon 1 specific primers were used to amplify the whole exon from the genomic DNA of patient 1, his mother and a control subject. The PCR products were then sent for sequencing, which showed that the parents were heterozygous for the insertion, while the patient was homozygous (figure 3.3). This insertion mutation causes a shift in the reading frame and results in a change of amino acid sequence from residue 24 onwards, and introduces a stop codon at position 31.

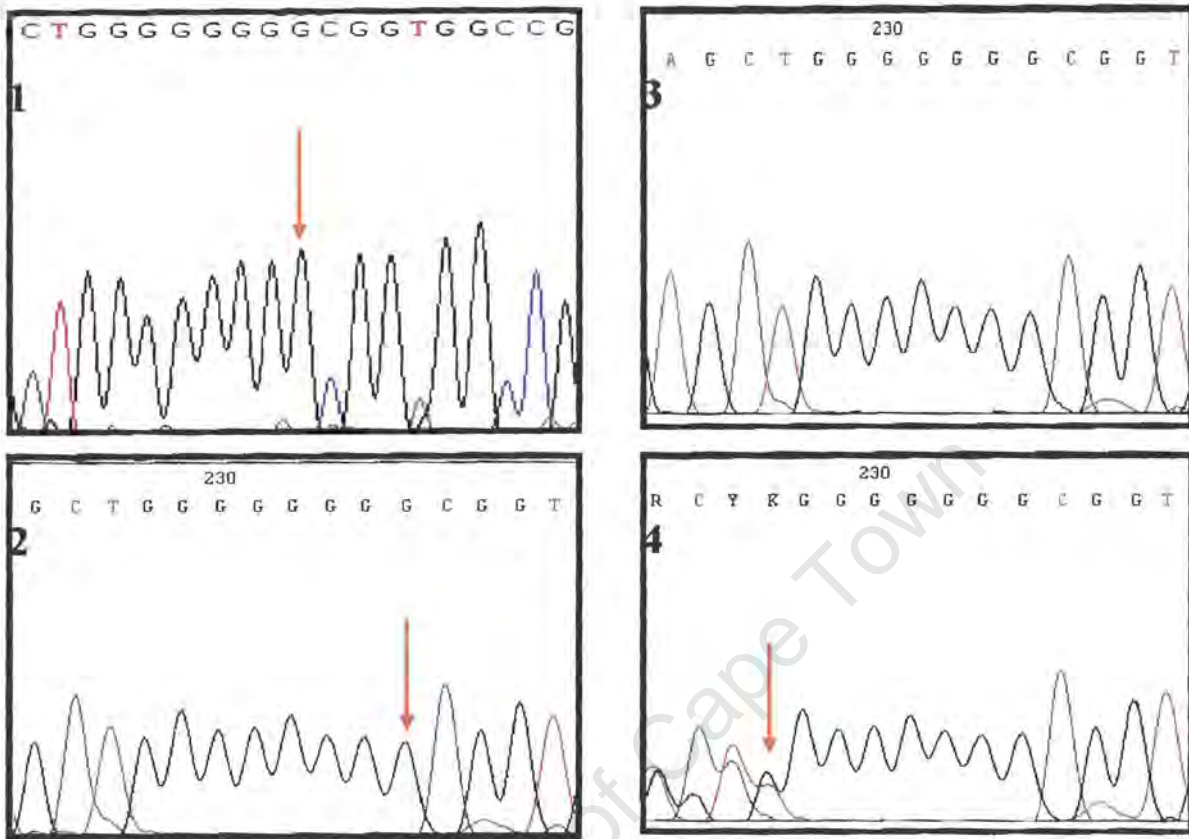


Fig. 3.3: Sequence traces of the PCR product spanning the run of 'G' residues from 65 to 71 of ARH cDNA. The complementary sequence of the reverse strand is given here. 1) Part of cDNA of ARH from patient 1 with 8 Gs 2) Part of exon 1 of ARH from the patient I with 8 Gs 3) Part of exon 1 of ARH from a control subject with 7 Gs 4) Part of exon 1 of ARH from the mother with a heterozygous pattern.

3.1.2.4 Mutational Analysis

The mutation was confirmed by *Bgl*I digestion of exon 1 PCR amplicons derived from the use of the mismatch primer pair, which introduces a *Bgl*I cutting site into the product from the normal *ARH* allele. Overnight digestion at 37°C thus gives two bands of 150 and 20bp for the product of the normal allele and an undigested product from the mutant allele. These data are shown in figure 3.4. Patient I is homozygous for the mutation and

shows only a non-restricted band, while his parents are clearly heterozygous with both the undigested 170bp and the 150bp restricted fragment (the 20bp band is not shown).

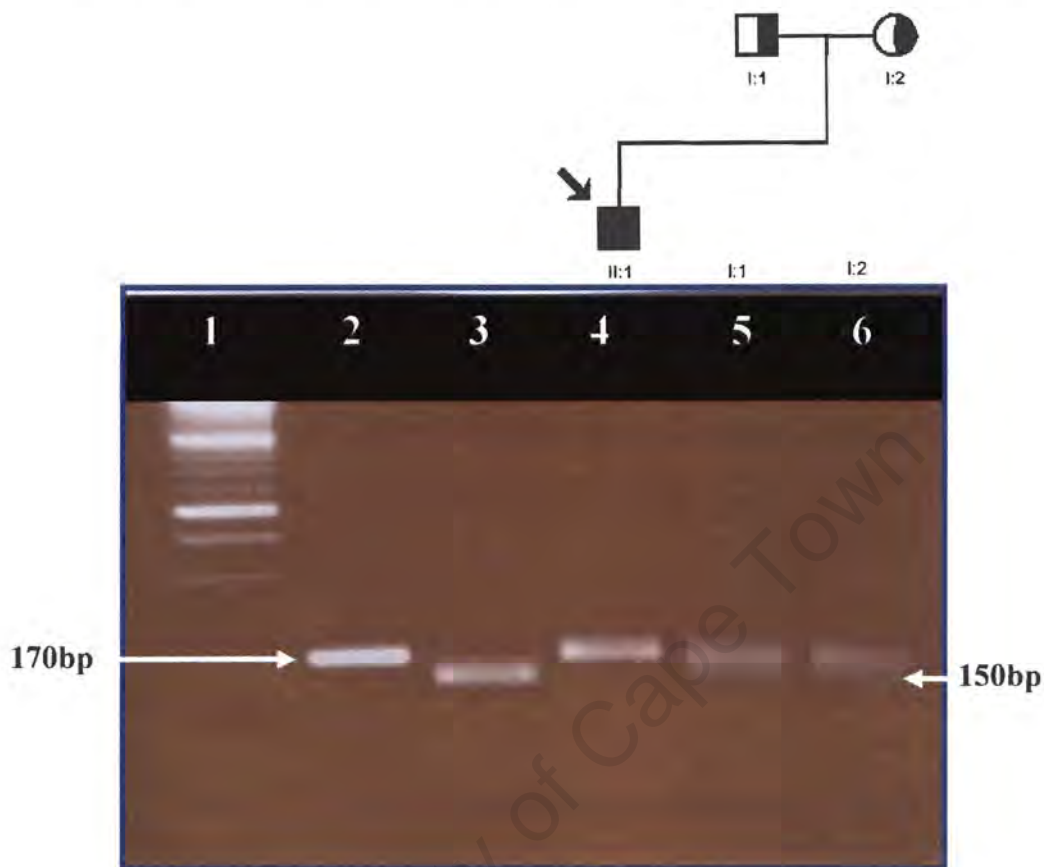


Fig. 3.4: Mutation detection by PCR and product restriction using *Bgl I*. Lane 1, Marker VIII (Roche); Lane 2, undigested control (170 bp); Lane 3, digested Control (150 bp); Lane 4, patient I (170 bp); Lane 5 and 6, father and mother respectively (170 and 150 bp). Digestion products were run on a 4% agarose gel

Some indication of the carrier frequency in black South Africans was sought by screening 400 cord-blood spots from a collection banked at the Red Cross Children's Hospital. These samples were amplified for exon 1 and the PCR product was digested with *Bgl I*. An agarose gel showing results from 10 of these cord-blood spots is given in figure 3.5, as an example. None of the 400 bloodspots tested positive for the insertion mutation.

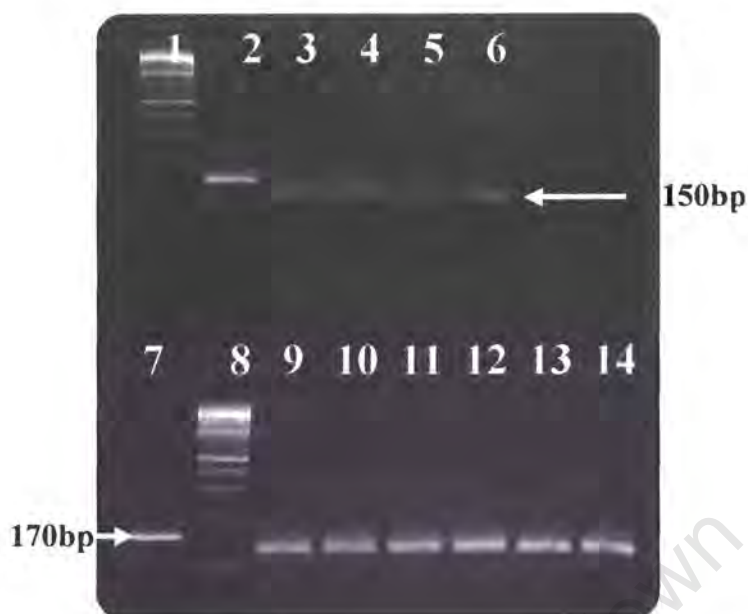


Fig. 3.5: A typical gel and fragmentation pattern observed in the screening of cord-blood spots for the insertion mutation. Shown are the digested PCR products from 10 blood spots run on 4% agarose. Lane 1 & 8, 100 bp DNA ladder; Lane 2 & 7, patient 1 (170bp; homozygous for the mutation) and Lanes 3-6 & 9-14, cord-bloods from black newborns (150 bp; the 20bp fragment is not visible in these gels).

3.1.2.5 Single Nucleotide Polymorphism (SNP)

Sequencing of patient 1 *ARH* cDNA, also revealed two single nucleotide polymorphisms, one in exon 6 and one in exon 7. These SNPs fall within the coding region and are of special interest in this study as they can be employed in haplotype analysis to define the genetic background of the mutant alleles in ARH patients.

3.1.2.5.1 Exon 6 SNP (C>T) NCBI ss# 35032447

Patient I was also found to be homozygous for a nucleotide change CCC→TCC at position 604 that codes for a serine instead of proline at amino acid position 202 of *ARH*. This SNP has been reported previously (Harada-Shiba et al. 2003) but had not been lodged in any of the SNP databases. This has now been done (NCBI ss# 35032447) with

the entry also containing details of the PCR assay developed for it in this study, as well as the frequency data in black South Africans.

This SNP was further detected by PCR amplification of a mismatch forward primer, to give a product of 120 bp, which on digestion with the restriction enzyme *Bsi* *YI* (CCN₇GG) gives two fragments of 102 and 18 bp for the 'T' allele. The digestion mix was loaded onto a 4% agarose gel, electrophoresed and the fragments visualized by UV.

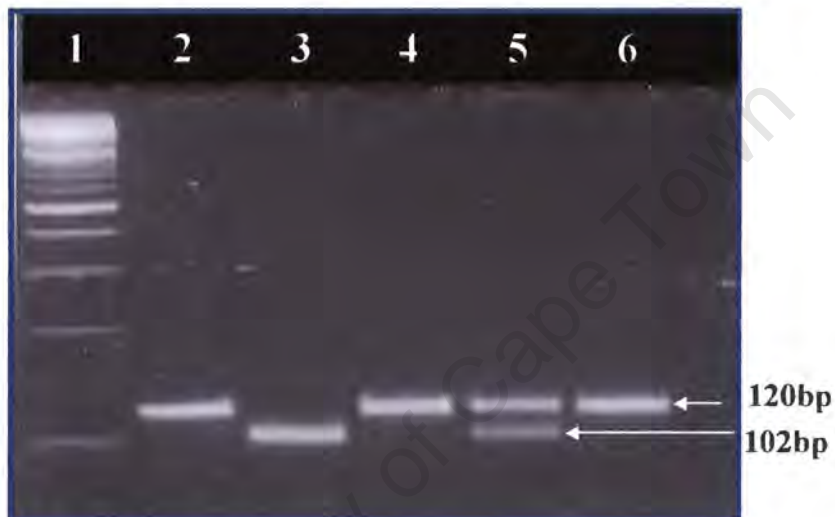


Fig. 3.6: Genotyping of the C→T SNP, in exon 6 of the ARH gene, by PCR and amplicon digestion with *Bsi* *YI*. Lane 1, Marker VIII (Roche); Lane 2, undigested control (120 bp); Lane 3, digested control (CC); Lane 4, patient I, digested (TT); Lane 5, paternal, digested (CT) and Lane 6, maternal, digested (TT). PCR products were run on a 4% agarose gel; the 12bp fragment is not visible.

The father of patient I is clearly heterozygous for the SNP (CT) (lane 5; Figure 3.6), while patient I and his mother are homozygous for the 'T' allele (TT). The 'T' allele is thus associated with the mutation in this family.

The frequency of this SNP was then determined by analyzing a 100 randomly selected cord blood spots from the Red Cross Children's Hospital blood spot collection.

Total alleles no analyzed = 200

Frequency of C=75/200 =0.375

T=125/200 =0.625

Hardy-Weinberg analysis

Subjection to Hardy-Weinberg was also determined by calculating the observed and expected frequency of the genotypes and by calculating the χ^2 .

	<u>CC</u>	<u>CT</u>	<u>TT</u>
Observed no	18	39	43
Expected no *	14	47	39

* $p = (18 \times 2 + 39) / 200 = 0.375$ (and hence $q = 1 - p = 0.625$)

Therefore frequency of CC = $p^2 = 0.375^2 = 0.14$

No of expected CC = $0.14 \times 100 = 14$

$\chi^2 = \% (\text{Observed} - \text{Expected})^2 / \text{Expected} = 2.822$ ($df=1$, $p= 0.0930$)

3.1.2.5.2 Exon 7 SNP (G→A) NCBI ss# 35032448

The second SNP found in patient I occurred in exon 7 at position 654 (ACG→ACA) and is silent, as it does not result in amino acid substitution. It does however, create an *Alu I* cutting site in the PCR product from the 'A' allele that yields two fragments 201 and 84bp (Figure 3.7). This SNP was neither published nor described in any SNP database and it was thus submitted to NCBI, where it was accorded the accession number 35032448.

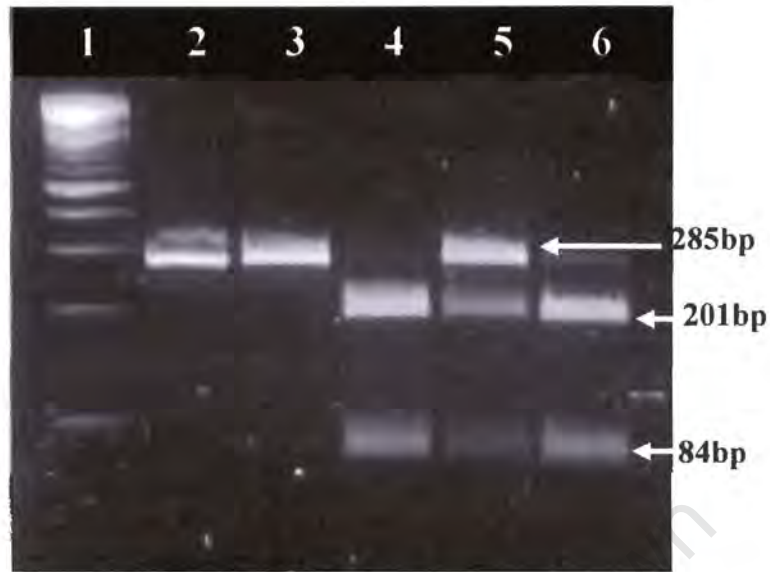


Fig. 3.7: Genotyping of the G → A SNP, in exon 7 of the ARH gene, by PCR and amplicon digestion with Alu I. Lane 1, Marker VIII (Roche); Lane 2, undigested control (214 bp); Lane 3, digested control (GG); Lane 4, digested Patient 1 (AA); Lane 5 and 6, digested paternal (GA) and maternal (AA), respectively. PCR products were run on a 4% agarose gel.

The band patterns in figure 3.7 show clearly that patient I and his mother are both homozygous for the ‘A’ allele, while his father is heterozygous. Frequency data was derived from genotyping 100 normal cord-bloods from the Red Cross Hospital cord-blood collection.

Total alleles no analyzed = 200

Frequency of G $163/200 = 0.815$

A $37/200 = 0.185$

Hardy-Weinberg analysis

	<u>GG</u>	<u>GA</u>	<u>AA</u>
Observed no	74	16	10
Expected no *	67.24	29.52	3.24

$$* p = (74 \times 2 + 16) / 200 = 0.82 \quad (\text{and hence } q = 1 - p = 0.18)$$

Therefore frequency of GG = $p^2 = 0.82^2 = 0.6724$

No of expected GG = $0.6724 \times 100 = 67.24$

$$\text{Chi}^2 = \% (\text{Observed} - \text{Expected})^2 / \text{Expected} = 20.150 \quad (df=1, p= 7.161 \text{ E-}06)$$

3.1.2.6 Haplotype Analysis

Haplotype analysis was carried out in order to contrast the genetic backgrounds of the South African 'G' insertion with that of the Iranian ARH patient, earlier reported with the same mutation. DNA from this patient was obtained as a kind gift from Prof John Kastelein, University of Amsterdam. The mutation was confirmed in the Iranian sample using the PCR assay described earlier (Section 2.3.5). These data are shown in Figure 3.8 and it can clearly be seen that the Iranian proband has the 'G' insertion, while his father is heterozygous.

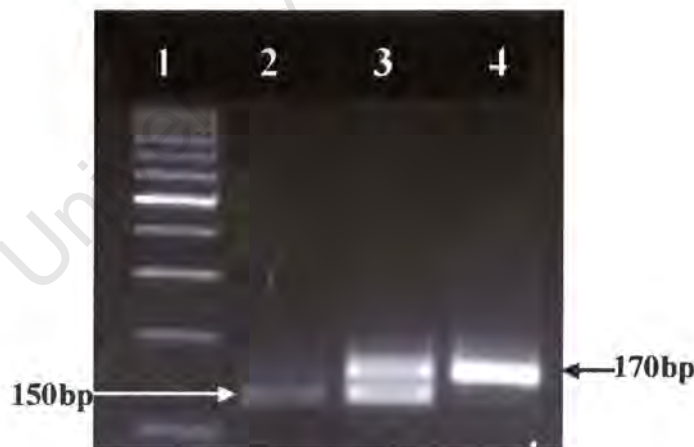


Fig. 3.8: Validation of the exon 1 'G' insertion mutation in the Iranian patient and his parents. PCR product digestion is carried out with Bgl I, which digests the product from the normal ARH allele only. Lane 1, 100 bp DNA ladder (Promega); Lane 2, control (150 bp); Lane 3, paternal (170 & 150); Lane 4, Iranian patient (170bp). Digestion products were run on a 4% agarose gel; the 20bp fragment is not shown.

To compare the genetic background of the African and Iranian mutations, haplotype analysis was done using the two exonic SNPs. Both exons were amplified and digested using their respective restriction enzymes. The genotyping data from the Iranian kindred are shown in figures 3.9 and 3.10. Haplotypes were constructed using mutation status and the two genotypes from the SNPs.

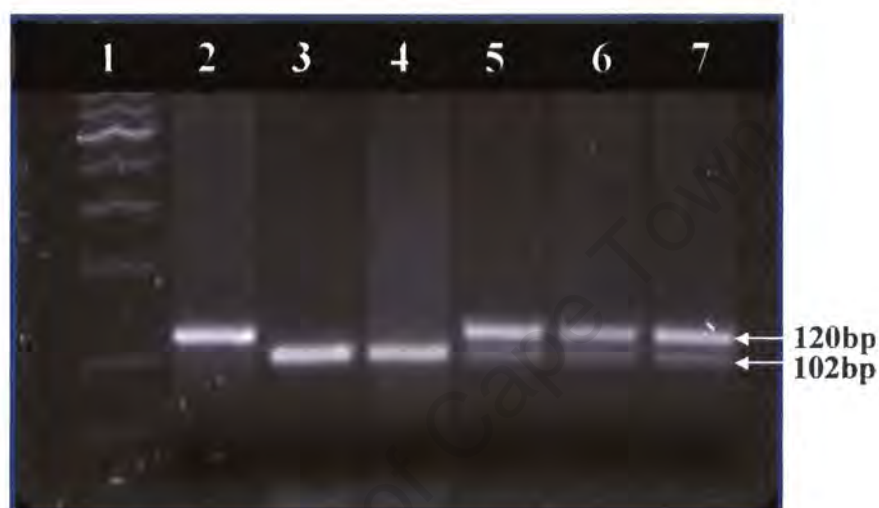


Fig. 3.9: Genotyping of the C \rightarrow T SNP, in exon 6 of the ARH gene, in the Iranian kindred. Digestion was carried out with the enzyme Bsi YI with the PCR digestion products being run on a 4% agarose gel. Lane 1, 100 bp DNA ladder; Lane 2, patient 1 (TT); Lane 3, digested control (CC); Lane 4, Iranian patient (CC); Lanes 5-7, father, mother and brother of the Iranian patient respectively (CT).

The 'G' insertion mutation in the Iranian patient is shown to associate with the 'C' allele of the exon 6 SNP.

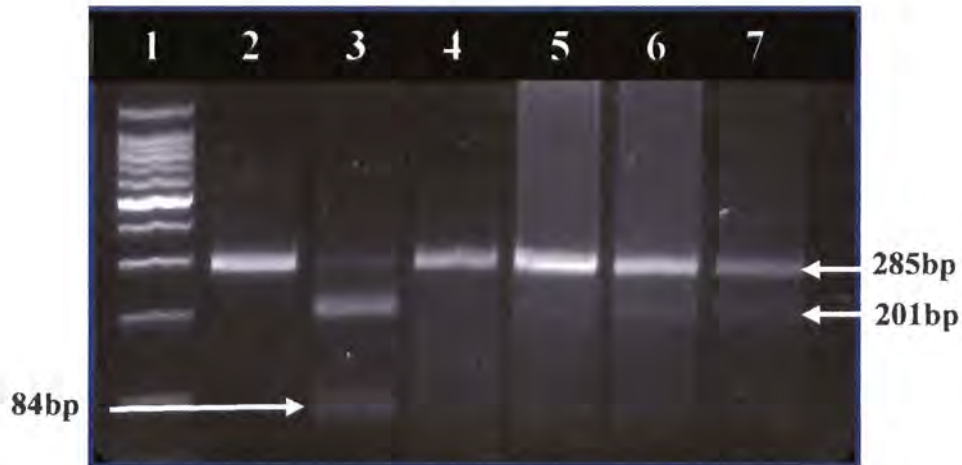


Fig. 3.10: Genotyping of the $G \rightarrow A$ SNP, in exon 7 of the ARH gene, in the Iranian kindred. Digestion was carried out with the enzyme Alu I and the digestion products were run on a 4% agarose gel. Lane 1, 100 bp DNA ladder (Promega); Lane 2, undigested control (285 bp); Lane 3, patient I (AA); Lane 4, Iranian patient (GG); Lanes 5-7, father, mother and brother of the Iranian patient (GA).

The 'G' insertion mutation in the Iranian patient is shown to associate with the 'G' allele of the exon 7 SNP.

Kindred I (patient I)

A partial family tree for this kindred is given in figure. 3.11. This tree only shows those individuals for whom DNA was available.

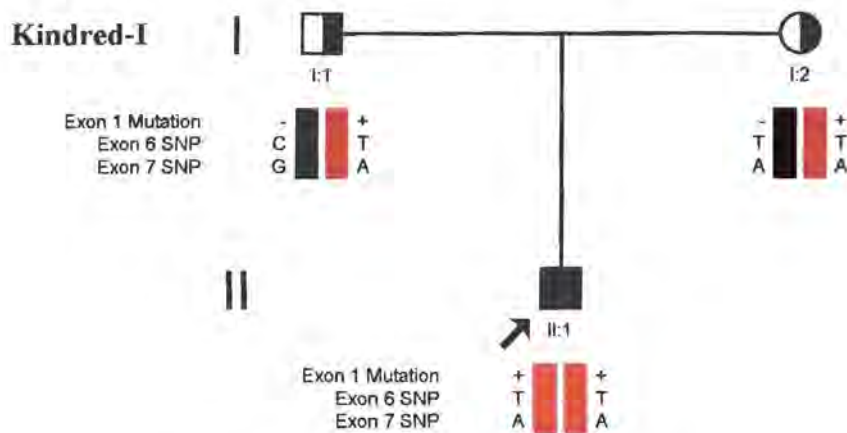


Fig. 3.11: A limited family tree of kindred-I showing the ARH gene haplotypes for the proband and his parents. The genotypes used to construct the haplotypes are listed on the left. The red bar marks the haplotype associated with the ARH mutation.

The haplotype characterizing the mutant allele in this family is (+, T, A).

Iranian Kindred

The proband in this kindred, was born of a consanguineous marriage. Haplotypes were constructed from both SNPs and it is clear that the mutant allele in this family is characterized by the (C, G) haplotype (Figure 3.12). Interestingly, the haplotype (T, A) that showed association with the disease in Kindred-I is associated with the normal locus in this family.

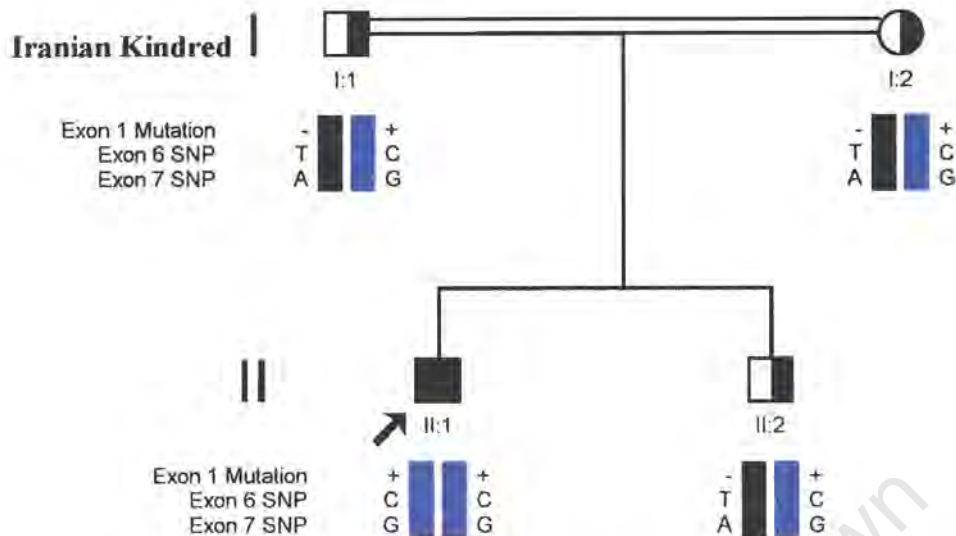


Fig. 3.12: The family tree of the Iranian kindred showing an arbitrary allocation of haplotypes. The genotypes used to construct the haplotypes are listed on the left.

3.2 PATIENT II (KINDRED II)

This patient, who was of coloured ancestry, also presented with clinical signs suggestive of homozygous FH. An indication that ARH may be the underlying defect came from the lipid profiles of his parents, which were slightly higher than normal but regarded as being too low for heterozygous FH (Table 2.1). It was decided to forgo LDLr functional studies in this patient and to first screen for the 'G' insertion mutation and determine the SNP haplotypes, and then to look directly for ARH synthesis in lymphocytes using the Western blotting technique.

3.2.1 DNA screening

DNA from patient II and his parents was screened for the 'G' insertion mutation as described previously (Section 2.3.5). All samples tested negatively for this mutation (Figure 3.13).



Fig. 3.13: Screening for the Exon 1 'G' insertion mutation in patient II and his parents. PCR products were digested with *Bgl* I and run on a 4% agarose gel. Lane 1, 100 bp DNA ladder (Promega); Lane 2, undigested control (170bp); Lane 3, digested control (150bp); Lane 4, patient I (170bp); Lane 5, patient II (150bp); Lanes 6 and 7, paternal and maternal amplicons respectively (150bp).

It is clear from the above gel that patient II does not have the 'G' insertion mutation as the products from both *ARH* alleles are fully digested, giving the normal 150 & 20bp banding pattern.

3.2.2 SNP analysis

Indirect support for an *ARH* defect in this patient was sought through haplotype analysis. In autosomal recessive disorders, patients are usually homozygous for SNPs found in the mutated gene because of consanguinity in the family or the inheritance of founder mutations. All family members were genotyped (figures 3.14 and 3.15) for the exon 6 and exon 7 SNPs, as previously described.

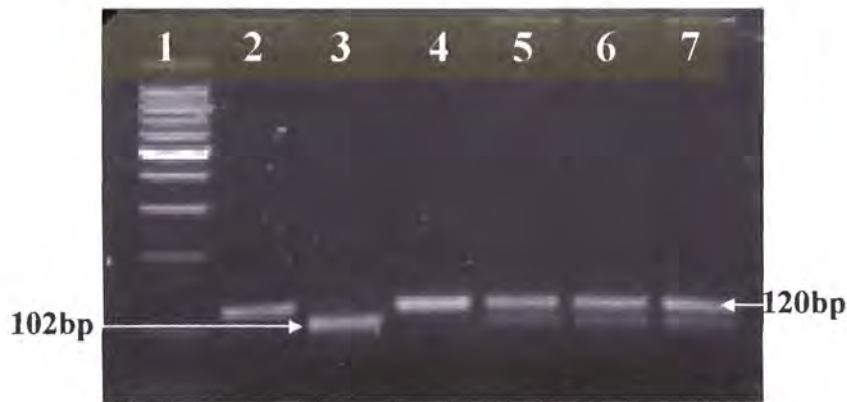


Fig. 3.14: Exon 6 SNP genotyping of DNA samples from kindred II. PCR products were digested with *Bsi* YI and run on a 4% agarose gel. Lane 1, 100 bp DNA ladder; Lane 2, undigested control (120 bp); Lane 3, digested control (CC); Lane 4, Patient I (TT); Lane 5, patient II (CT); Lanes 6 & 7, paternal and maternal amplicons respectively (CT).

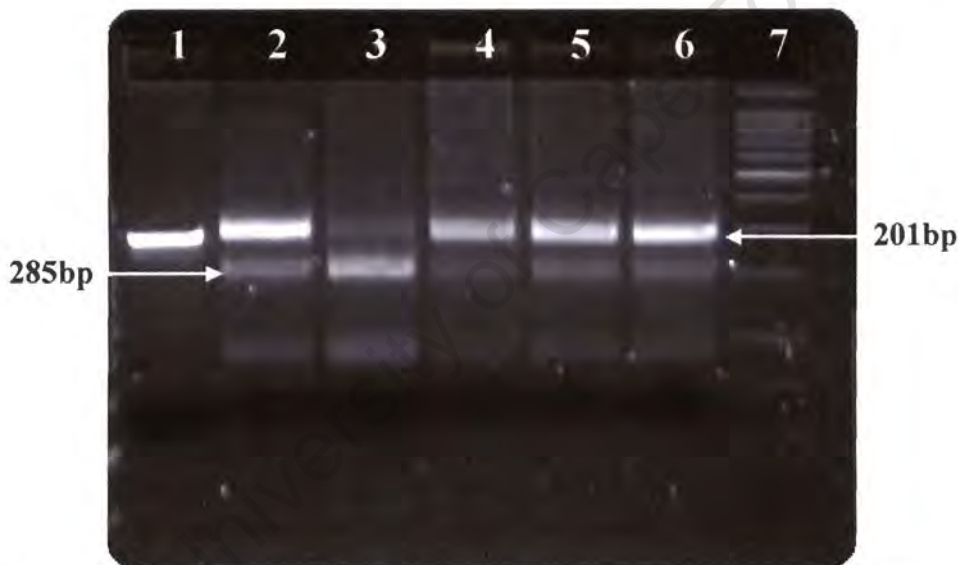


Fig. 3.15: Exon 7 SNP genotyping of DNA samples from kindred II. PCR products were digested with *Alu* I and run on a 4% agarose gel. Lane 1, undigested control DNA (285 bp); Lane 2, heterozygous control (GA); Lane 3, patient I (AA); Lane 4, patient II (GA); Lanes 5 & 6, paternal and maternal amplicons, respectively (GA); Lane 7, 100bp DNA ladder.

Patient II and his parents were found to be heterozygous for the two SNPs (CT and GA for exon 6 and exon 7 respectively; figure 3.14 and 3.15) giving four possible haplotypes for each member. An arbitrary allocation of haplotypes is shown in figure 3.16. What

ever the true allocation of haplotypes is, it is clear that patient II is not homozygous for these two genetic markers at the *ARH* locus.

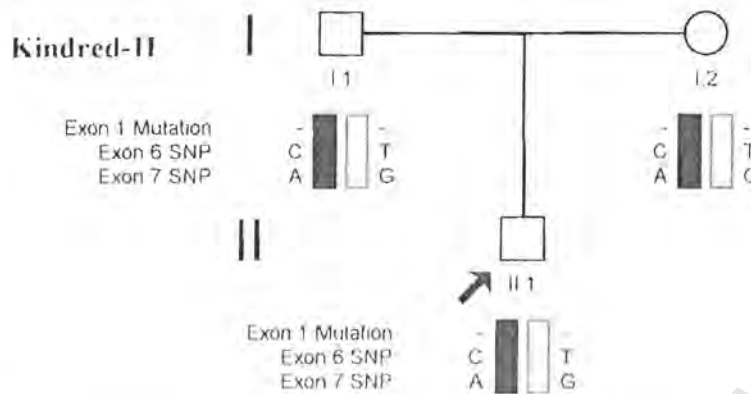


Fig. 3.16 The family tree of kindred-II showing an arbitrary allocation of haplotypes. The genotypes used to construct the haplotypes are listed on the left.

3.2.3 Western Blotting

Western blotting was used to look for the presence of ARH in transformed lymphocytes from patient II and to confirm the predicted absence of ARH in cultured fibroblasts from patient I, control fibroblasts and transformed lymphocytes were run in parallel; the protocol followed is described in detail in the methods section. Shown in figure 3.17(A) is the nitrocellulose membrane containing the proteins transferred from the polyacrylamide gel. The proteins have been stained with Ponceau stain and it is clear that all isolates contain cellular proteins, and in similar amounts. Figure 3.17(B) shows the resulting Western blot and clearly indicates the presence of ARH in the lymphocyte extract of patient II, while confirming the absence of ARH in the fibroblast extract from patient I.

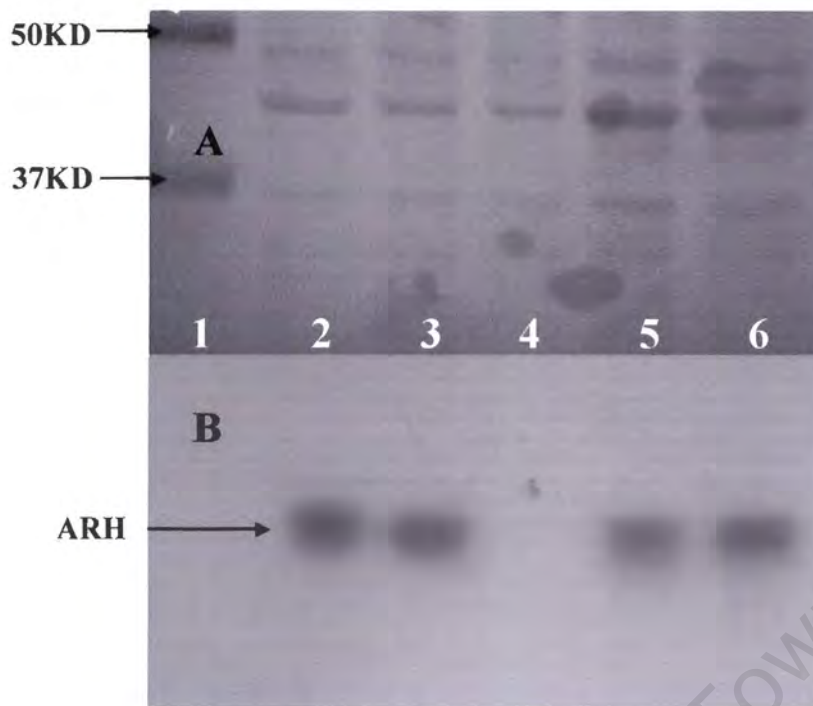


Fig. 3.17: Immunoblot analysis of ARH protein in cultured fibroblasts and transformed lymphocytes. A) A Ponceau stain of the nitrocellulose membrane post transfer. B) A Western blot probed with anti-ARH antibody. Lane 1, Protein standards (BIO-RAD); Lane 2, control fibroblasts; Lane 3, fibroblasts from FH homozygote; Lane 4, fibroblasts of patient I; Lane 5, lymphoblasts, father of Patient II; Lane 6, lymphoblasts, patient II.

DISCUSSION & CONCLUSION

Familial hypercholesterolemia (FH) is one of the most common and severe monogenic lipoprotein disorders. It is characterized by vascular accumulation of LDL cholesterol, premature atherosclerosis and coronary artery disease. This autosomal dominant disorder is caused either by a mutation in the receptor of LDL (LDLr) or in the receptor-binding domain of its ligand, apoB-100. Over the years, rare cases of FH with autosomal recessive inheritance have been described. These patients have a clinical phenotype of homozygous FH with almost normal LDLr activity in cultured fibroblasts while showing defective LDL uptake in EBV transformed lymphocytes and macrophages (Soutar et al. 2003). The gene associated with the disorder was discovered by typical mapping studies in affected kindreds and was named after the disorder, autosomal recessive hypercholesterolemia (*ARH*) gene. The protein, ARH, encoded from this gene facilitates the endocytic machinery of the coated pit by interaction with the cytoplasmic tail of LDLr, AP-2 and clathrin.

The aims of this study were two fold, firstly to confirm the provisional diagnosis of ARH deficiency in two kindreds, where the probands presented with clinical manifestations of FH, but with family histories and parental lipid levels not suggestive of FH, and secondly, to set up the molecular and cellular methodologies required to detect and characterize genetic defects in ARH biology.

Kindred I, Patient I

FH and ARH can be distinguished at the cellular level by LDL binding and uptake studies on cultured skin fibroblasts or transformed lymphocytes, where the former shows defective function in both cell types while the latter manifests a defect in lymphocyte

function only. Attempts to make this distinction in cells from patient I, were unsuccessful as his lymphocytes repeatedly failed to survive after transformation. This phenomenon has been reported by other laboratories and would indicate some critical function of ARH in transformed lymphocytes. The comments from other laboratories are based on their experience of ARH deficient patients and on transfection and mutagenesis investigations into ARH function (unpublished data Prof Linton Traub, University of Pittsburgh School of Medicine, Pittsburgh).

LDL binding and uptake studies on cultured fibroblasts from patient I showed an average of 99.4% LDLr activity when compared to control fibroblast cultures (Table 3.1). This finding ruled against a diagnosis of homozygous FH as these patients have <30% internalisation of labeled LDL when compared to controls.

Further molecular analysis to elucidate the genetic background of the disorder revealed an insertion of a guanine base (G), in a string of seven consecutive 'G' nucleotides (bp 65 to 71) in exon 1 of *ARH*. Runs of bases are considered as hot spots for mutations as they are susceptible to insertion and deletion by DNA slippage. This slippage is caused by strand displacement of the nascent DNA strand, followed by an out-of-register pairing (Levinson & Gutman 1987). This is evident in *ARH*, as two out of the 13 mutations reported are found in runs of bases, one each in the run of 'G's in exon 1 and 'C's in exon 6 (Garcia et al. 2001; Harada-Shiba et al. 2003).

This 'G' insertion introduces an early premature stop codon at position 33 after encoding 24 amino acids of random sequence. The predicted truncated ARH protein lacks the entire PTB domain and other functional segments and is thus incompatible with normal ARH function. This defect is further compounded by the likelihood of mRNA instability

due to nonsense-mediated mRNA decay (NMD), a well-described phenomenon that permits the elimination of mutant mRNAs from their translatable pools (Wilund et al. 2002). The fact that we were able to amplify cDNA from this patient does not exclude NMD, as any residual mRNA would yield to the sensitivity of PCR.

Final confirmation of ARH deficiency in patient I was obtained from Western blotting, where fibroblast protein extracts failed to yield a signal when probed with antibodies to ARH; significant ~37-kDa ARH protein band were visualized in both control extracts (Figure 3.17).

The paucity of ARH patients described in South Africa and worldwide would suggest that the background carrier frequency is very low, compounded by the fact that ARH is a structural protein; these molecules are typically more tolerant of amino acid replacements than enzymes, for example. Some indication of the 'G' insertion mutation carrier frequency in black South Africans was sought by screening DNA from 400 newborns. These newborns were all of Xhosa descent and while our patient was of Zulu ancestry, they both stem from the Nguni tribes that settled in South Africa. None of the newborn samples tested positive for the insertion mutation, so the real frequency is likely to be very low; this single screen would suggest that the frequency is $<1/400$.

Turning to the second objective, the frequencies of both SNPs found in the black patient were also analyzed. The SNP in exon 6 substitutes proline for serine with a frequency of 0.37 and 0.63 for the 'C' and 'T' alleles respectively, in black South Africans. The expected and observed genotype frequencies for this SNP were similar and were subjected to Hardy-Weinberg analysis ($\chi^2=2.822$ $df=1$, $p=0.0930$), which is not a significant change). The polymorphism in exon 7 (G→A) was silent and novel. The

alleles 'G' and 'A' have a frequency of 0.82 and 0.18 respectively, in black South Africans. There were fewer observed heterozygous (16%) for this SNP than expected (30%) and when genotypes were subjected to Hardy-Weinberg analysis gave $\chi^2=20.150$, which is a significant change ($df=1$, $p= 7.161 \text{ E-}06$). The reason for this SNP being out of Hardy-Weinberg equilibrium is most likely an indication of technical shortcoming e.g. incomplete digestion leading to the misclassification of individuals.

Another question to be addressed here was whether the South African 'G' insertion mutation was from the same mutational event that led to ARH deficiency in the Iranian patient in whom a 'G' insertion, in the identical run of 'G's, was first described. Haplotype analysis using two intragenic SNPs was employed to address this issue. Haplotypes in the kindreds were constructed from the genotypes derived from the SNPs and their mutation status (+, -). Both patients were homozygous for both SNPs allowing unequivocal assignment of the (T, A) and (C, G) haplotypes to the mutant alleles in the South African and Iranian patients, respectively. This divergence of genotypes and haplotypes is highly suggestive of two mutational events occurring on different genetic backgrounds, presumably at two different points in time. It is also impossible to determine where the inserted G was located as there are eight likely sites for the run of seven Gs.

Kindred II, Patient II

The second kindred to be investigated for suspected ARH deficiency was of mixed ancestry and it was considered expedient to screen for the insertion mutation detected in our black proband; DNA diagnosis in galactosemia and cystic fibrosis have shown that

gene mutations, common in the black population are often found in the communities of mixed ancestry (Henderson et al. 2002; Goldman et al. 2001).

DNA was amplified using the in-house PCR assay developed to detect this mutation, with the patient testing negative for the insertion mutation. This finding thus slightly decreased the likelihood of ARH deficiency in this kindred.

Genotyping of the exon 6 and exon 7 SNPs was also carried out. This was done on the understanding that should the provisional diagnosis of ARH deficiency be correct, then homozygosity for these genotypes would most likely be recorded. This has been the finding in most autosomal recessive conditions where founder gene inheritance and consanguinity predominate. Heterozygosity for both SNPs was found, decreasing further the likelihood of ARH deficiency.

Good evidence against a primary ARH defect was gained from the Western blotting investigations where a strong ~37kDa ARH signal was found in protein extracts from transformed lymphocytes of patient II. This finding, although very significant needs to be tempered by the fact that Western blotting makes little comment on the function of proteins. The rare situation of an intact but non-functional ARH remains a possibility and needs further investigation by LDL uptake studies in transformed lymphoblasts. It is possible that some other protein involved in ARH function is defective. Of interest is that four truncation mutants of ARH have now been examined by Western blotting, and none of them have detected ARH mass in cell extracts. The only mutant to show ARH mass, has a splice-site defect that produces an internal deletion of 26 amino acids and an ARH product that is of lower molecular weight (Wilund et al. 2002; Arca et al. 2002; Harada-Shiba et al. 2003; Canizales-Quinteros et al. 2005).

The most likely diagnosis in this kindred is one of FH, with the atypical finding being the low LDLc levels in the heterozygous parents. It is remotely conceivable that these atypical levels may be the result of parental heterozygosity for FH in conjunction with heterozygosity for proprotein convertase subtilisin/kexin type 9 serine protease (PCSK9), a newly discovered cholesterol lowering factor. Loss-of-function mutations in the PCSK9 gene are associated with about 40% reduction in plasma levels of LDL cholesterol. The mutations probably reduce plasma LDLc by increasing the number of LDL-receptors (Cohen et al. 2005). FH heterozygous patients with PCSK9 loss-of-function mutations might therefore, have normal or slightly elevated cholesterol profile due to the increase in the number of LDL-receptors. These combined mutations might explain the low LDLc levels found in the parents of patient II. Further studies will be required to elucidate the genetic cause of this disorder as this research project only focused on ARH defects.

In conclusion, this study describes the first black kindred affected by ARH in South Africa. The black patient has an insertion mutation in exon 1 of *ARH*. This mutation results in a severely truncated ARH protein, which is most likely rapidly degraded and explains the absence of ARH signaling on Western blotting. A PCR and restriction enzyme based assay has been developed which will allow for the screening of other patients and selected cohorts. Identification of this rare ARH genetic defect in the black patient has also provided a new experience in treatment of hypercholesterolemia and allows for appropriate genetic counseling for the family.

APPENDIX A:

BUFFERS AND SOLUTIONS

TISSUE CULTURE

Lysis Buffer

Tris	50mM
CaCl ₂	2mM
NaCl	80mM
Triton	1% V/V (add protease inhibitor cocktail before use)

Mycoplasma Detection

Fluorescent stain

Hoechst No 33258	5mg
Hanks balanced salt solution	100ml

Working solution

Hanks' balanced salt solution 1×

Mounting fluid

Citric acid	0.1M
Na ₂ HPO ₄ ·2H ₂ O	0.2M
Glycerol	50ml (make up to 100ml pH=5.5)

Fixative solution

Glacial acetic acid-methanol 1:3

Protein Concentration (Markwell Modification of Lowry)

Solution A

Na ₂ CO ₃ (anhydrous)	2%
NaOH	0.4%
Na ₂ Tartrate	0.16%
SDS	1% (to a final volume of 1000ml)

Solution B

CuSO ₄ ·5H ₂ O	4% (dissolve to a final volume of 100ml)
--------------------------------------	--

Solution C

Solution A	100parts
Solution B	1part (made up immediately before use)

MOLECULAR ASSAYS

Agarose Gel-loading Buffer I

bromophenol blue	0.25% w/v
xylene cyanol FF	0.25%
sucrose in water	40% (w/v)

Agarose Gel tracking dye

bromophenol blue	0.25 % w/v
sucrose	40% w/v
EDTA	20mM (pH 8.0)

TAE electrophoresis buffer (50X 1litre)

Tris	242g
Glacial acetic acid	57.1ml
EDTA (0.5M)	100ml
Adjust volume with distilled water	

Tris-EDTA (TE) buffer:

Tris	10mM
EDTA	1mM
Adjust to pH 7.6 with HCL	

Western Blotting

Acrylamide (100ml; 30%)

Acrylamide	150g
Methylenbisaacrylamid	4g
Na dodecyl sulph	0.5g (Adjust volume to a final volume of 500ml and store in a dark bottle at 4°C)

Ammonium Persulfate (AMP) (0.5ml; 10%)

Dissolve 0.05g in 0.5ml of distilled water (prepare freshly, prior to gel mix)

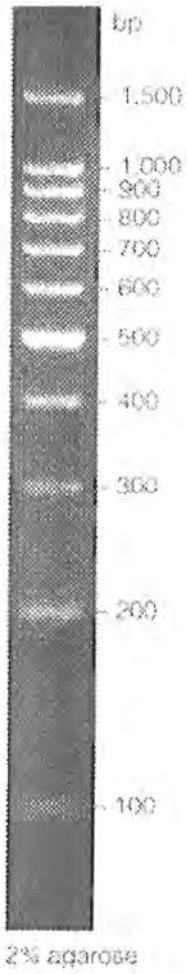
Blocking/blotting buffer

Tween 20 in PBS	0.05% (make from package)
-----------------	---------------------------

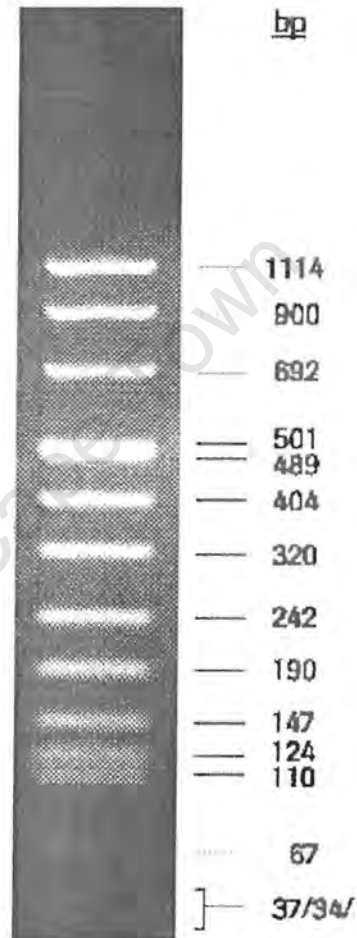
Dry milk	5% (need about 100 ml/membrane)
Running buffer	
Tris	25mM
Glycine	192mM
SDS	0.1%
SDS Gel Buffer	
Tris	68.100g
Sodium dodecylsulph	0.5g (Adjust pH to 8.8 and bring it 500ml)
SDS PAGE Gel mix (10%)	
Acrylamide stock (30%)	10ml
Gel buffer	10ml
SDS	0.1%
Ammonium Persulfate (10%)	400µl
TEMED	16µl
Stack gel mix	
Acrylamide stock (30%)	0.93ml
Stack buffer	6.07ml
Ammonium Persulfate (10%)	112µl
TEMED	5µl
Stripping buffer	
SDS	2%
Tris-base	62.5mM
B-mercaptoethanol	100mM (add right before using)
Transfer Buffer	
Tris	0.025M
Glycine	0.192M
	Bring volume to 4 L and store at 4°C. Take 800 ml of this and add 200 ml of methanol when doing transfer
Washing Buffers:	
Tween 20	0.05% in PBS

APPENDIX B:

MOLECULAR WEIGHT MARKERS



100 bp DNA ladder 100-1,500 (Promega)



DNA molecular weight Marker VIII (ROCHE)

REFERENCES:

- Al-Kateb H, Bahring S, Hoffmann K, Strauch K, Busjahn A, Nurnberg G, Jouma M, et al (2002) Mutation in the ARH gene and a chromosome 13q locus influence cholesterol levels in a new form of digenic-recessive familial hypercholesterolemia. *Circ Res* 90:951-958
- Arca M, Zuliani G, Wilund K, Campagna F, Fellin R, Bertolini S, Calandra S, et al (2002) Autosomal recessive hypercholesterolaemia in Sardinia, Italy, and mutations in *ARH*: a clinical and molecular genetic analysis. *Lancet* 359:841-847
- Barbagallo CM, Emmanuele G, Cefalu AB, Fiore B, Noto D, Mazzarino MC, Pace A, et al (2003) Autosomal recessive hypercholesterolemia in a Sicilian kindred harboring the 432insA mutation of the ARH gene. *Atherosclerosis* 166:395-400
- Broeders N, Knoop C, Abramowicz D (1999) Drug treatment of lipid disorders. *N Engl J Med* 341:2020-1
- Brown MS, Goldstein JL (1986) A receptor-mediated pathway for cholesterol homeostasis. *Science* 232:34-47
- Canizales-Quinteros S, Aguilar-Salinas CA, Huertas-Vazquez A, Ordonez-Sanchez ML, Rodriguez-Torres M, Venturas-Gallegos JL, Riba L, et al (2005) A novel ARH splice site mutation in a Mexican kindred with autosomal recessive hypercholesterolemia. *Hum Genet* 116:114-120
- Chen GC, Hardman DA, Hamilton RL, Mendel CM, Schilling JW, Zhu S, Lau K, Wong JS, Kane JP (1989) Distribution of lipid-binding regions in human apolipoprotein B-100. *Biochemistry* 28:2477-84

Chen TR (1977) In situ detection of mycoplasma contamination in cell cultures by fluorescent Hoechst 33258 stain. *Exp Cell Res* 104:255-62

Chen WJ, Goldstein JL, Brown MS (1990) NPXY, a sequence often found in cytoplasmic tails, is required for coated pit-mediated internalisation of the low density lipoprotein receptor. *J Biol Chem* 265:3116-3123

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162:156-159

Chung BH, Wilkinson T, Geer JC, Segrest JP (1980) Preparative and quantitative isolation of plasma lipoproteins: rapid, single discontinuous density gradient ultracentrifugation in a vertical rotor. *J Lipid Res* 21:284-91

Ciccarese M, Pacifico A, Tonolo G, Pintus P, Nikoshkov A, Zuliani G, Fellin R, et al (2000) A new locus for autosomal recessive hypercholesterolemia maps to human chromosome 15q25-q26. *Am J Hum Genet* 66:453-460

Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH (2005) Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet* 2005 37:161-5

Eden ER, Patel DD, Sun XM, Burden JJ, Themis M, Edwards M, Lee P, et al (2002) Restoration of LDL receptor function in cells from patients with autosomal recessive hypercholesterolemia by retroviral expression of ARH1. *J Clin Invest* 110:1695-1702

Eden ER, Naoumova RP, Burden JJ, McCarthy MI, Soutar AK (2001) Use of homozygosity mapping to identify a region on chromosome 1 bearing a defective gene that causes autosomal recessive homozygous hypercholesterolemia in two unrelated families. *Am J Hum Genet* 68:653-660

Fellin R, Zuliani G, Arca M, Pintus P, Pacifico A, Montali A, Corsini A, et al (2003) Clinical and biochemical characterisation of patients with autosomal recessive hypercholesterolemia (ARH). *Nutr Metab Cardiovasc Dis* 13:278-86

Frolov A, Zielinski SE, Crowley JR, Dudley-Rucker N, Schaffer JE, Ory DS (2003) NPC1 and NPC2 regulate cellular cholesterol homeostasis through generation of low density lipoprotein cholesterol-derived oxysterols. *J Biol Chem* 278:25517-25

Garcia CK, Wilund K, Arca M, Zuliani G, Fellin R, Maioli M, Calandra S, et al (2001) Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. *Science* 292:1394-1398

Goldman A, Labrum R, Claustres M, Desgeorges M, Guittard C, Wallace A, Ramsay M (2001) The molecular basis of cystic fibrosis in South Africa. *Clin Genet* 59:37-41.

Goldstein JL, Brown MS (1974) Binding and degradation of low density lipoproteins by cultured human fibroblasts. Comparison of cells from a normal subject and from a patient with homozygous familial hypercholesterolemia. *J Biol Chem* 249:5153-62

Harada-Shiba M, Takagi A, Marutsuka K, Moriguchi S, Yagyu H, Ishibashi S, Asada Y, Yokoyama S (2004) Disruption of autosomal recessive hypercholesterolemia gene shows different phenotype in vitro and in vivo. *Circ Res* 95:945-52

Harada-Shiba M, Takagi A, Miyamoto Y, Tsushima M, Ikeda Y, Yokoyama S, Yamamoto A (2003) Clinical features and genetic analysis of autosomal recessive hypercholesterolemia. *J Clin Endocrinol Metab* 88:2541-2547

He G, Gupta S, Yi M, Michaely P, Hobbs HH, Cohen JC (2002) ARH is a modular adaptor protein that interacts with the LDL receptor, clathrin, and AP-2 *J Biol Chem* 277:44044-44049

Henderson H, Leisegang F, Brown R, Eley B (2002) The clinical and molecular spectrum of galactosemia in patients from the Cape Town region of South Africa. *BMC Pediatr* 2:7

Hubacek JA, Hyatt T (2004) ARH missense polymorphisms and plasma cholesterol levels. *Clin Chem Lab Med* 42:989-90

John E. Coligan, Ada M. Kruisbeek, David H. Margulies, Ethan M. Shevach, Warren Strober (eds), (Series Editor: Richard Coico) (1991) *Current Protocols in Immunology* Unit 7.22

Jones C, Hammer RE, Li WP, Cohen JC, Hobbs HH, Herz J (2003) Normal sorting but defective endocytosis of the low density lipoprotein receptor in mice with autosomal recessive hypercholesterolemia. *J Biol Chem* 278:29024-30

Khachadurian AK (1964) The inheritance of essential familial hypercholesterolemia. *Am J Med* 37:402-7

Khachadurian AK, Uthman SM (1973) Experiences with the homozygous cases of familial hypercholesterolemia: A report of 52 patients. *Nutr Metab* 15:132-140

Kibbey RG, Rizo J, Gierasch LM, Anderson RG (1998) The LDL receptor clustering motif interacts with the clathrin terminal domain in a reverse turn conformation. *J Biol Chem* 273:59-67

Kotze MJ, De Villiers WJ, Steyn K, Kriek JA, Marais AD, Langenhoven E, Herbert JS, et al (1993) Phenotypic variation among familial hypercholesterolemics heterozygous for either one of two Afrikaner founder LDL receptor mutations. *Arterioscler Thromb* 13:1460-8

Knoblauch H, Muller-Myhsok B, Busjahn A, Ben Avi L, Bähring S, Baron H, Heath SC, et al (2000) A cholesterol-lowering gene maps to chromosome 13q. *Am J Hum Genet* 66:157-66.

Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol* 4:203-21.

Lind S, Olsson AG, Eriksson M, Rudling M, Eggertsen G, Angelin B (2004) Autosomal recessive hypercholesterolaemia: normalization of plasma LDL cholesterol by ezetimibe in combination with statin treatment. *J Intern Med* 256:406-12

Mahley RW, Ji ZS (1995) Remnant lipoprotein metabolism: key pathways involving cell-surface heparan sulfate proteoglycans and apolipoprotein E. *J Lipid Res* 40:1-16

Marais AD, Firth JC, Blom DJ (2004) Homozygous familial hypercholesterolemia and its management. *Semin Vasc Med* 4:43-50

Markwell, M.A.K., S.M. Haas, L.L. Bieber, and N.E. Tolbert (1978) A modification of the Lowry procedure to simplify protein determination in the membrane and lipoprotein samples. *Anal. Biochem* 87:206-210

Michaely P, Li WP, Anderson RG, Cohen JC, Hobbs HH (2004) The modular adaptor protein ARH is required for low density lipoprotein (LDL) binding and internalization but not for LDL receptor clustering in coated pits. *J Biol Chem* 279:34023-31

Mishra SK, Keyel PA, Edeling MA, Dupin AL, Owen DJ, Traub LM (2005) Functional dissection of an AP-2 beta 2 appendage-binding sequence within the autosomal recessive hypercholesterolemia (ARH) protein. *J Biol Chem* 2005 Feb 22

Mishra SK, Keyel PA, Hawryluk MJ, Agostinelli NR, Watkins SC, Traub LM (2002) Disabled-2 exhibits the properties of a cargo-selective endocytic clathrin adaptor. *EMBO J* 21:4915-4926

Mishra SK, Watkins SC, Traub LM (2002) The autosomal recessive hypercholesterolemia (ARH) protein interfaces directly with the clathrin-coat machinery. *Proc Natl Acad Sci USA* 99:16099-16104

Norman D, Sun XM, Bourbon M, Knight BL, Naoumova RP, Soutar AK (1999) Characterization of a novel cellular defect in patients with phenotypic homozygous familial hypercholesterolemia. *J Clin Invest* 104:619-628

Noviello C, Vito P, Lopez P, Abdallah M, D'Adami L (2003) Autosomal recessive hypercholesterolemia protein interacts with and regulates the cell surface level of Alzheimer's Amyloid 2 precursor protein. *J Biol Chem* 278:31843-7

Paul J. (1975) *Cell and Tissue Culture*, 5th ed. New York: Churchill Livingstone.

Pitas RE, Innerarity TL, Weinstein JN, Mahley RW (1981) Acetoacetylated lipoproteins used to distinguish fibroblasts from macrophages in vitro by fluorescence microscopy. *Arteriosclerosis* 1:177-85

Rader DJ, Cohen J, Hobbs HH (2003) Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. *J Clin Invest* 111:1795-1803

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* (2nd edition), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Scriver CH, Beaudet AL, Sly SW, Valle D (1995) *The Metabolic and Molecular Bases of Inherited Disease* (7th edition). pp 1994-1995 & 1849-1850 Times Roman Composition Company, Inc.

Sly WS, Grubb J (1979) Isolation of fibroblasts from patients. *Methods Enzymol* 58:444-50

Soutar AK, Naoumova RP, Traub LM (2003) Genetics, clinical phenotype, and molecular cell biology of autosomal recessive hypercholesterolemia. *Arterioscler Thromb Vasc Biol* 23:1963-70

ter Haar E, Harrison SC, Kirchhausen T (2000) Peptide-in-groove interactions link target proteins to the beta-propeller of clathrin. *Proc Natl Acad Sci USA* 97:1096-100

van der Westhuyzen DR, Coetzee GA, Demasius IP, Harley EH, Gevers W, Baker SG, Seftel HC (1984) Low density lipoprotein receptor mutations in South African homozygous familial hypercholesterolemic patients. *Arteriosclerosis* 4:238-47

Wilund KR, Yi M, Campagna F, Arca M, Zuliani G, Fellin R, Ho YK, et al (2002) Molecular mechanisms of autosomal recessive hypercholesterolemia. *Hum Mol Genet* 11:3019-3030

Zuliani G, Arca M, Signore A, Bader G, Fazio S, Chianelli M, Bellosta S, et al (1999) Characterization of a new form of inherited hypercholesterolemia: familial recessive hypercholesterolemia *Arterioscler Thromb Vasc Biol* 19:802-809