Autosomal Recessive Hypercholesterolemia Caused by Mutations in a Putative LDL Receptor Adaptor Protein

Christine Kim Garcia,1 Kenneth Wilund,1,2 Marcello Arca,3 Giovanni Zuliani,4 Renato Fellini,4 Mario Maioli,5 Sebastiano Calandra,6 Stefano Bertolini,7 Fausto Cossu,8 Nick Grishin,9 Robert Barnes,1 Jonathan C. Cohen,1 Helen H. Hobbs1,2

Atherogenic low density lipoproteins are cleared from the circulation by hepatic low density lipoprotein receptors (LDLR). Two inherited forms of hypercholesterolemia result from loss of LDLR activity: autosomal dominant familial hypercholesterolemia (FH), caused by mutations in the LDLR gene, and autosomal recessive hypercholesterolemia (ARH), of unknown etiology. Here we map the ARH locus to a ~1-centimorgan interval on chromosome 1p35 and identify six mutations in a gene encoding a putative adaptor protein (ARH). ARH appears to have a tissue-specific role in LDLR function, as it is required in liver but not in fibroblasts.

The liver is the major site of synthesis and clearance of cholesterol ester–rich lipoproteins. More than 70% of circulating LDL is removed from the blood via hepatic LDLR-mediated endocytosis. In individuals with two mutant LDLR alleles (homozygous FH),...
the rate of clearance of LDL from the blood is
decreased, resulting in hypercholesterolemia,
xanthomatosis (deposition of cholesterol in
skin and tendons) and premature coronary
artery disease (CAD) (1). LDLR activity in
cultured skin fibroblasts from FH homozy-
gotes is also very low (7). A rare autosomal
recessive form of hypercholesterolemia
(ARH) that clinically resembles FH but is not
due to mutations in LDLR has been described
(2–9). These patients have markedly im-
paired hepatic LDLR function but normal or
only modestly reduced LDLR function in
cultured fibroblasts (3–5, 7, 8).

To elucidate the molecular basis of ARH
we performed a whole-genome linkage study
in four ARH families (Fig. 1A). Two were of
Sardinian origin and had low LDL clearance
rates in vivo (ARH1 and ARH2) (8), and two
were of Lebanese origin (ARH3 and ARH4),
including the original family described with
this disorder (2). The probands of the four
families were offspring of consanguineous
unions and all families showed horizontal
transmission of hypercholesterolemia. All af-
fected family members were severely hyper-
cholesterolemic and had very large xantho-
mas; some had premature CAD (Table 1).
Plasma LDL levels tended to be lower and the
onset of symptomatic CAD somewhat
later in these probands than in FH homozy-
gotes. LDLR function in cultured fibroblasts
from affected family members were normal
or only moderately reduced (3, 7, 8), thus
ruling out a diagnosis of homozygous FH.

Multipoint linkage analysis revealed sig-
nificant linkage [logarithm of the odds ratio
for linkage (lod) score, 7.4] to a 5.7-cM
interval on 1p35, demarcated by the polymor-
phic loci D1S2864 and D1S2787 (Fig. 1B)
(10). This interval overlaps to a chromosomal
region on 1p35–p36 linked to ARH in two
other families (11). We found no linkage to
15q25–q26, which was previously found to
be associated with ARH in five Sardinian

Fig. 1. Pedigrees (A), linkage analysis (B),
and fine mapping (C) of the ARH gene. (A)
The four pedigrees used for gene map-
ning (ARH1 to ARH4) are shown. ARH1 and
ARH2 are Sardinians, and ARH3 and ARH4
are Lebanese. Fasting plasma total choles-
terol levels (TC) are shown. (B) Distribu-
tion of lod scores in the linked region on
chromosome 1. A total genome scan was
performed initially in ARH1 and ARH2 and
then additional markers were typed in all
four families. The maximum lod score was
7.4 over a 1-cM region on chromosome 1. (C) Fine map-
ning within the linked region in ARH2.
Genomic DNA was
extracted from whole
blood that had been
collected from the
deceased probands
and stored at −20°C
for more than 10
years, or from fresh
leukocytes isolated
from venous blood.
The region of ho-
mozygosity shared by
the affected individu-
als in this family is
boxed. Squares, males;
circles, females;
double lines,
consanguineous mat-
ings; filled squares,
affected individuals.
We refined the linked region to a ~1-cM interval extending from D1S1152 to D1S2885 by identifying a region of homozygosity shared by all affected family members but not by a normocholesterolemic sibling in ARH2 (Fig. 1C) (12). The coding sequences of 13 genes that mapped to this interval were screened for sequence variation by PCR and the single strand conformation polymorphism (SSCP) technique (13). Two abnormally migrating bands were identified in the predicted coding sequences of a cDNA (DKFZp586D0624) in probands from ARH1 and ARH3. The gene structure and predicted amino acid sequence of the protein encoded by this cDNA are shown in Fig. 2. The gene spans ~25 kb and has nine exons and eight introns (Fig. 2A). The predicted amino acid sequence contains a 170–amino acid motif that shares considerable sequence similarity with the phosphotyrosine binding (PTB) domains of many adaptor proteins (14, 15) (Fig. 2B). PTB domains bind the consensus sequence NPXY, which is present in the cytoplasmic domains of several cell-surface receptors, including the epidermal growth factor receptor (16), the insulin receptor (17), nerve growth factor receptor (TrkA) (18), and the LDLR (19, 20). The integrity of the NPXY sequence in the cytoplasmic tail of the LDLR is absolutely required for internalization (19, 20), and the LDLR has been shown in vitro to bind other proteins containing PTB domains (21, 22).

**Fig. 2.** Gene structure (A), predicted amino acid sequence (B), and location of mutations in ARH probands (C). (A) The cDNA for DKFZp586D0624 (GenBank Accession Number AL117654) was prepared by PCR amplification of reverse-transcribed liver poly(A)⁺ mRNA. A PstI clone containing the entire gene (290N7, Incyte Genomics, Inc., Palo Alto) was used to amplify the introns and to sequence intron-exon boundaries. ARH is encoded by nine exons and spans ~25 kb. Filled rectangles, exons; lines, introns. (B) The predicted amino acid sequences of human, mouse, and Xenopus ARH. Numbers to the right correspond to human sequence. The alignment of the inferred amino acid sequences displays 67% identity among the three proteins. The regions of amino acid sequence identity are boxed. ARH has a highly conserved PTB domain at the amino terminus (indicated by blue shading, 89% identity). Alignment was constructed with PSI-BLAST (23). The boundaries of the PTB domain are according to Pfam 6.0 database [[35, 36]; domain PF00640]. (C) Schematic representation of ARH showing the location of the mutations identified in this study.

**Table 1.** Molecular defects in ARH and clinical characteristics of probands in four families with ARH (Fig. 1). Genomic DNA was extracted from cultured fibroblasts or leukocytes. The coding regions of the gene were screened for sequence variation using SSCP and dyeoxy sequencing. The nucleotides and amino acids were numbered from the A of the initiation codon (ATG). The age at the time of diagnosis is provided. The plasma cholesterol and LDL-cholesterol levels were measured by the referring physician. LDLR activity was assessed as described in the reference and is provided as a percentage of normal control fibroblasts studied simultaneously. Abbreviations: TC, fasting plasma total cholesterol; ref, reference; ins, insertion; F, female; CAD, symptomatic or documented coronary artery disease; AS, aortic murmur or stenosis; yr, years; M, male; ND, not done; NIDDM, non-insulin-dependent diabetes; Tx, treatment; MI, myocardial infarction; del, deletion; X, stop; amino acids: W, tryptophan; Q, glutamine; P, proline; H, histidine.

<table>
<thead>
<tr>
<th>Family</th>
<th>Nucleotide change</th>
<th>Amino acid change</th>
<th>Origin</th>
<th>Age/sex</th>
<th>Plasma TC/LDL-C (mg/dl)</th>
<th>LDLR activity in fibroblasts</th>
<th>Comments</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARH1</td>
<td>c.432insA</td>
<td>170Stop</td>
<td>Nuoro, Sardinia</td>
<td>20/F</td>
<td>530/460</td>
<td>–70</td>
<td>CAD, AS; eight relatives died at &lt;33 years</td>
<td>(3, 7–9)</td>
</tr>
<tr>
<td>ARH2</td>
<td>c.65G&gt;A</td>
<td>W22X</td>
<td>Olbia, Sardinia</td>
<td>23/M</td>
<td>540/464</td>
<td>–80</td>
<td>CAD, AS</td>
<td>(8)</td>
</tr>
<tr>
<td>ARH3</td>
<td>c.406C&gt;T</td>
<td>Q136X</td>
<td>Beirut, Lebanon</td>
<td>21/M</td>
<td>440/ND</td>
<td>60–70</td>
<td>Father, NIDDM</td>
<td>(2, 3)</td>
</tr>
<tr>
<td></td>
<td>c.605C&gt;A</td>
<td>P202H</td>
<td>Lebanon</td>
<td>17/F</td>
<td>610/520</td>
<td>100</td>
<td>Father, MI, 28 years</td>
<td>(3)</td>
</tr>
<tr>
<td>ARH5</td>
<td>c.72insG</td>
<td>33Stop</td>
<td>Iran</td>
<td>10/M</td>
<td>637/598</td>
<td>ND</td>
<td>AS, TC = 321 on Tx</td>
<td>(3)</td>
</tr>
<tr>
<td>ARH6</td>
<td>c.71delG</td>
<td>55Stop</td>
<td>USA</td>
<td>15/F</td>
<td>800</td>
<td>ND</td>
<td>AS, TC = 321 on Tx</td>
<td>(3)</td>
</tr>
</tbody>
</table>
Database searches (23) revealed orthologous proteins in mouse and Xenopus that share 89% sequence identity with the human protein in the PTB domain (Fig. 2B). Several regions in the COOH-terminal half of these protein are also highly conserved. These blocks do not appear to be shared with other proteins currently in the database. The closest paralogues of ARH are the Drosophila NUMB protein (24) and the Caenorhabditis elegans CED-6, an adaptor protein involved in cell engulfment (25). These proteins share 33% (52%) and 34% (60%) sequence identity (similarity) with the human protein, respectively.

The coding region of ARH1 was sequenced using genomic DNA from the affected family members of ARH1, ARH2, ARH3, and ARH4 (Fig. 1). Two different mutations that cause premature termination of translation were identified in the Sardinian families. The affected individuals in ARH1 were homozygous for a single base-pair insertion in exon 4 (Table 1) that introduces a premature termination codon (Fig. 1). Two different mutations that cause premature termination of translation were identified in ARH4 (26, 27). Both mutations account for ARH in these 12 apparently unrelated Sardinian probands probably reflects genetic drift, which has been observed for other diseases on the island (26, 27). There was significant overlap in the distribution of the mutation on the island; neither mutation was found in 50 normolipidemic Sardinians (28).

The four affected Lebanese siblings in ARH3 (Fig. 1) were homozygous for a nonsense mutation in exon 136, which stops translation in the terminal region of the PTB binding domain. Both ARH4 probands were homozygous for a missense mutation substituting a histidine for proline at amino acid 202, which is outside the PTB domain. The LDL activity was normal in the ARH3 fibroblasts (3). Neither of the mutations found in ARH3 or ARH4 was present in 15 normolipidemic individuals from Lebanon, seven unrelated Lebanese FH homoyzoytes with a molecularly defined defect in the LDLR gene (29), or in 50 normocholesterolemic Caucasians. ARH probands from two other unrelated consanguineous families (ARH5 and ARH6) were homozygous for different frameshift mutations located in a string of seven guanine residues in exon 1 (Table 1). Both mutations are predicted to truncate the protein near the NH2-terminus (Fig. 2C).

We performed Northern blot analysis to assess the size and relative abundance of the ARH mRNA in cultured fibroblasts from the probands of the ARH1, ARH3, and ARH4 families (Fig. 3A). A 3.1-kb mRNA was detected in the control fibroblasts. In contrast to the LDLR mRNA, the levels of ARH mRNA were not affected by the addition of sterols to the medium. Only trace amounts of ARH mRNA were detected in the ARH1 fibroblasts, in which both ARH alleles contained the frameshift mutation in exon 4, and in the ARH3 fibroblasts, in which were homozygous for a nonsense mutation in exon 4. Normal levels of ARH mRNA were present in the ARH4 fibroblasts, which harbored a homozygous missense mutation. Studies are in progress to determine the functional effects of this mutation. ARH4 was the only family in which the parents had evidence of a possible defect in cholesterol metabolism; the father had a myocardial infarction at age 28. In one ARH family described by Norman et al. (6), both parents had moderately elevated plasma LDL-cholesterol levels. These observations raise the possibility that some ARH mutations result in codominant, rather than recessive hypercholesterolemia.

Preliminary results using a mammalian two-hybrid system indicate that the PTB domain of ARH interacts with the cytoplasmic tail of the LDLR, and additional studies are under way to characterize the specificity and physiological significance of this interaction. PTB domains differ in their selectivity for NPXY sequences in different proteins, which allows for specificity in the biological response (30). For example, the Drosophila SHC (30) and mouse Disabled (21) adaptor proteins bind to only a subset of NPXY sequences. ARH appears to be a close phenocopy of homozygous FH, which suggests that all clinical sequelae of ARH mutations are attributable to defective LDLR activity, and this in turn suggests that ARH binds specifically to the LDLR. However, although both ARH and LDLR appear to be nearly ubiquitously expressed (Fig. 3B), LDLR expression is relatively low in some of the same tissues that express high levels of ARH (kidney, placenta) (1), raising the possibility that this protein may be involved in other receptor pathways. None of the 16 probands examined in this study have other obvious shared phenotypes that would suggest defective signaling or functioning of NPXY-containing proteins, with the possible exception of NIDDM (non-insulin-dependent diabetes mellitus) (Table 1).

The defect in LDLR function in ARH appears to be not only receptor-specific, but also tissue-specific. We have been unable to identify a consistent defect in LDLR function (binding, uptake or internalization) in cultured fibroblasts from ARH patients. It is possible that another PTB domain protein compensates for the absence of ARH in cultured fibroblasts, or that adaptor molecules are not required for receptor-mediated endocytosis of LDL in fibroblasts and possibly other extrahepatic cells.

ARH may participate in a step in the itinerary of the LDLR pathway that is specific to polarized cells like hepatocytes (31). ARH may be required for trafficking of LDLR to the basolateral surface. In contrast to fibroblasts, LDLRs in hepatocytes do not cluster in coated pits; conceivably, ARH may target the LDLR to the coated pit after the receptor binds LDL (32). Alternatively, ARH may participate in the recycling of the LDLR from the lysosome to the basolateral cell surface after dissociating from LDL. Although the specific role of ARH in the functioning of the LDLR remains to be defined, the crucial role of this protein is revealed by the profound hypercholesterolemia that occurs in this disease.

References and Notes
3. J. L. Goldstein, M. S. Brown, unpublished observations.
10. A whole-genome linkage analysis was performed using 450 polymorphic DNA markers at ~8-cM intervals (Cooperative Human Linkage Center/Weber Human Screening Set Version B, Research Genetics, Inc., Huntsville). Markers were genotyped in selected family members from ARH1 (IV.1, IV.2, V.1, V.2, V.3) and ARH2 (IV.1, IV.2, V.1, V.2, V.3) (Fig. 1). Linkage analysis using GENEHUNTER (33) and CRIMAP (34), ruled out linkage to 93% of the genome. An additional 70 genetic markers covering the 14 genomic regions that could not be excluded on the initial genomewide screen were genotyped in all numbered members of the four families (Fig. 1A). Linkage to a region on 1p35 was found with a lod score of 7.4. The affected siblings in ARH1 and ARH3 had inherited alleles identical by descent in this region but were not homozygous for any of the markers. The two siblings of ARH4 shared a 44-cM region of homozygosity in this region.


12. The centrometric boundary of homozygosity was defined by D1S2885 (family members IV.3 and IV.6 in ARH2), which is telomeric to GGA2D04 on the physical map (www.ncbi.nlm.nih.gov). The telomeric boundary was delineated by marker D1S1152 in individual IV.4, who was normolipemic and yet was homozygous for the markers distal to D1S1152. The exact position of D1S1152 was determined through genetic analysis of crossovers in ARH2 and CEPH family no. 1362. The coding regions of the genes located in the physical region between markers D1S1152 and D1S2885 were screened for sequence variations.


19. C. G. Davis et al., Cell 45, 15 (1986).


38. We thank E. Kern, Y. Liao, R. Wilson, J. Hunter, and B. Crider and especially T. Hyatt for excellent technical assistance; A. Cao, A. Montali, F. Campagna, and N. Kalaany for assistance in obtaining human blood samples; M. Brown, J. Goldstein, and D. W. Russell for helpful discussions; J. Kastelein and J. C. Defesche for providing genomic DNA from an Iranian ARH patient; and A. K. Khachadurian, M. J. Chapman, J. L. DeGennes, and J. M. Hoeg for providing skin fibroblasts from ARH patients and M. Brown and J. Goldstein for making them available to us. Supported by POI-408, the W. M. Keck Foundation and the Donald W. Reynolds Cardiovascular Clinical Research Center (H-J+H), Telethon grant E. 0565 (M.A.), and NIH training grant HL07360 (K.W.).

Include this information when citing this paper.