Molecular basis of familial hypercholesterolaemia from structure of LDL receptor module

Deborah Fass*†, Stephen Blacklow*†‡, Peter S. Kim*† & James M. Berger†

* *Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA*
† *Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA*
‡ *Present address: Department of Pathology, Stanford University, 300 Pasteur Drive, Stanford, California 94305-5324, USA.*

The low-density lipoprotein receptor (LDLR) is responsible for the uptake of cholesterol-containing lipoprotein particles into cells[1,2]. The amino-terminal region of LDLR, which consists of seven tandemly repeated, ~40-amino-acid, cysteine-rich modules (LDL-A modules), mediates binding to lipoproteins[3,4]. LDL-A modules are biologically ubiquitous domains, found in over 100 proteins in the sequence database[5]. The structure of ligand-binding repeat 5 (LR5) of the LDLR, determined to 1.7 Å resolution by X-ray crystallography and presented here, contains a calcium ion coordinated by acidic residues that lie at the carboxy-terminal end of the domain and are conserved among LDL-A modules. Naturally occurring point mutations found in patients with the disease familial hypercholesterolaemia[6] alter residues that directly coordinate Ca^{2+} or that serve as scaffolding residues of LR5.

Familial hypercholesterolaemia (FH) is a genetic disease caused by defects in LDLR function[6]. Normally the LDL-A modules of the LDLR mediate binding to the lipoproteins LDL and β-VLDL[3,4], which are then imported into the cell by receptor-mediated endocytosis[1,2]. Mutations of the LDLR that affect this process result in failure to clear lipoproteins from the circulation, pathologically elevated blood cholesterol and premature heart disease[6]. Many point mutations that cause FH map to the fifth LDL-A module of the LDLR (LR5), particularly to a cluster of acidic residues near the carboxy-terminal end of the module (Fig. 1).

To investigate the molecular basis of FH, we determined the structure of LR5 by X-ray crystallography (Fig. 2). Crystals of LR5 were grown from ammonium sulphate with sucrose as a cryoprotectant (see Methods). Reflection data were collected at −180 °C and

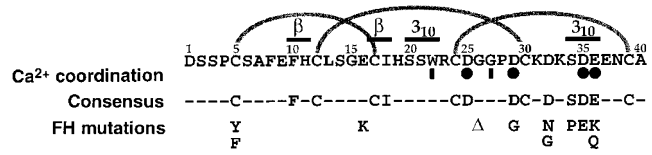

Figure 1 Sequence of LR5. Numbering refers to residue position in the LR5 peptide[9] and corresponds to that used in the text. Above the sequence, disulphide connectivity is indicated by loops, and locations of secondary structure elements are shown by black bars. Below the sequence, black circles highlight residues with side chains that participate in Ca^{2+} coordination. Vertical bars indicate residues with backbone carbonyl oxygens that contribute to Ca^{2+} coordination. Residues explicitly shown in the consensus sequence are conserved in at least 6 of the 7 amino-terminal LDLR modules[24]. The positions and identities of mutations causing FH[6] are displayed under the consensus sequence. The triangle indicates deletion of a residue.

phases were determined by multiple isomorphous replacement (MIR) with heavy-atom compounds (Table 1). The working and free *R*-factors to 1.7 Å resolution are 20.9 and 24.0%, respectively, with good protein geometry (Table 1).

The striking feature of the LR5 structure is that it is organized around a calcium ion. Although the overall topology of LR5 is similar to that of homologous modules 1 and 2 determined by NMR[7,8], the NMR structures do not show evidence of Ca^{2+} binding. The residues in LR5 that coordinate the Ca^{2+} ion through their side chains are Asp 25, Asp 29, Asp 35 and Glu 36 (Figs 1, 3), the same acidic residues that are conserved among the LDL-A modules. The backbone carbonyls of Trp 22 and Gly 27 complete the octahedral, or six-ligand, coordination geometry (Fig. 3). The ion binding site explains the Ca^{2+} requirement for proper folding and disulphide bond formation of LR5 (ref. 9) and provides a rationale for the Ca^{2+} requirement in lipoprotein binding[10].

The effects of FH mutations that occur in module 5 of the LDLR (Fig. 1) can be explained by the LR5 structure. These mutations fall into two categories: elimination or misplacement of direct Ca^{2+} ligands, and removal of hydrogen bonds or disulphide bonds that help maintain the backbone topology. In the first category, mutation of Asp 29, Asp 35 or Glu 36 destroys the Ca^{2+}-binding site. Deletion of Gly 26 alters the spacing between the Ca^{2+} coordinating residues. In the second category, Glu 16 is in a position to form a hydrogen bond with the backbone amides of residues 16 and 31 by using both of its carboxylate oxygens, as well as with the Ser 14 side-chain hydroxyl group. This hydrogen-bonding network of Glu 16 is apparently crucial for stabilizing the tertiary structure of the module. Similarly, the Asp 32 side chain makes a hydrogen bond

Table 1 Structure determination and refinement

Derivative	Resolution (Å)	Completeness (%)	R_{sym}*	R_{iso}†	No. of sites	Phasing power‡	Cullis *R*§	Cullis $R_{anomalous}$
Native	1.7	94.3	0.054	–	–	–	–	–
K_2PtCl_4	2.0	97.9	0.089	0.279	4	2.25	0.58	0.96
OsO_4	2.4	90.0	0.034	0.204	1	1.16	0.81	–
$AuCl_3$	2.0	99.9	0.066	0.188	1	1.00	0.86	–

Overall figure of merit 0.530

Number of non-hydrogen atoms			Resolution (Å)	No. of reflections working/free	R_{cryst}/R_{free}	R.m.s. deviations bonds/angles
Prot	Wat	Ion				
279	30	1 Ca^{2+} 1 SO_4^{2-}	15.0–1.7	2732/227 (7.67%)	0.209/0.240	0.013 Å/2.498°

The LR5 model that was built into the MIR electron density map contains residues 4–40 of the domain (Fig. 1); residues 1–3 were not visible in the MIR map or in $2F_o - F_c$ maps calculated with refined models of residues 4–40. The model also contains one highly coordinated calcium ion, 30 water molecules and one sulphate ion.
* $R_{sym} = \Sigma\Sigma_j |I_j - \langle I \rangle| / \Sigma I_j$, where I_j is the intensity measurement for reflection *j* and $\langle I \rangle$ is the mean intensity for multiply recorded and symmetry-related reflections.
† $R_{iso} = \Sigma \big| |F_{ph}| - |F_p| \big| / \Sigma |F_p|$, where F_{ph} and F_p are the derivative and native structure factors, respectively.
‡ Phasing power $= \langle F_h \rangle / E$, where $\langle F_h \rangle$ is the root-mean-square heavy-atom structure factor and *E* is the residual lack of closure error.
§ Cullis $R = \Sigma \big| |F_{ph} \pm F_p| - |F_{h,c}| \big| / \Sigma |F_{ph} \pm F_p|$ where $F_{h,c}$ is the calculated heavy-atom structure factor.
‖ Cullis $R_{anomalous} = \Sigma \big| |F_{ph}(+) - F_{ph}(-)| - 2F''_h \sin\alpha_p \big| / \Sigma |F_{ph}(+) - F_{ph}(-)|$, where α_p is the calculated protein phase.
$R_{cryst,free} = \Sigma \big| |F_{obs}| - |F_{calc}| \big| / \Sigma |F_{obs}|$, where the crystallographic and free *R* factors are calculated using the working and free reflection sets, respectively.

with the backbone amide of residue 14 and forms an Asx turn[11] by hydrogen bonding to the backbone amide of the conserved Ser at position 34. The use of both carboxylate oxygens explains why the relatively conservative Asp-to-Asn mutation at this position results in FH. Finally, Ser 34 is apparently conserved in LDL-A modules to act as a hydrogen-bond acceptor from the backbone amide of residue 13, and as a donor to the side chain of Asp 32. A proline at this position, in addition to kinking the chain near Ca^{2+} ligands, lacks side-chain hydrogen-bonding potential and the backbone amide to participate in the Asx turn with Asp 32. Not surprisingly, this mutation results in FH.

In addition to justifying the effects of FH mutations, the structure of LR5 calls into question existing models for normal LDLR function. These models postulate that the interaction between LDLR and its target occurs between the conserved acidic motif in the LDLR ligand-binding modules and positively charged residues

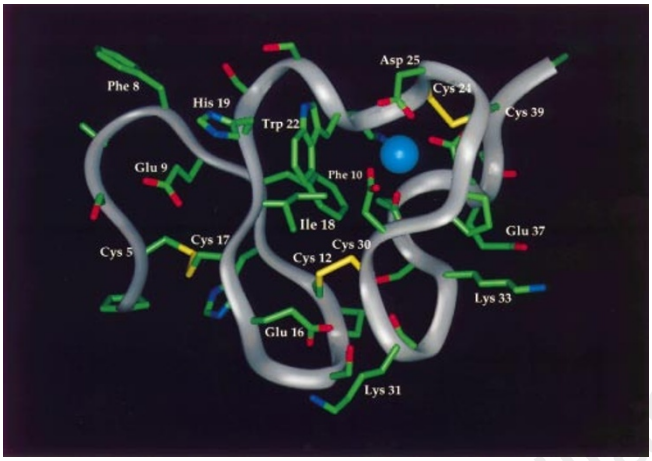

Figure 2 Structure of LR5. Side chains are superimposed on a ribbon trace of the module backbone. Cysteine and additional residues are numbered as reference points. The Ca^{2+} ion is indicated by a blue sphere. FH mutations are located at structurally significant residues (see text). The figure was generated using Insight II (Biosym Technologies). The LR5 structure has uniformly low *B*-factors for all residues after residue 4 (average main-chain *B*-factor = 10.7 Å^2; average side-chain *B*-factor = 11.0 Å^2 for residues 5–40).

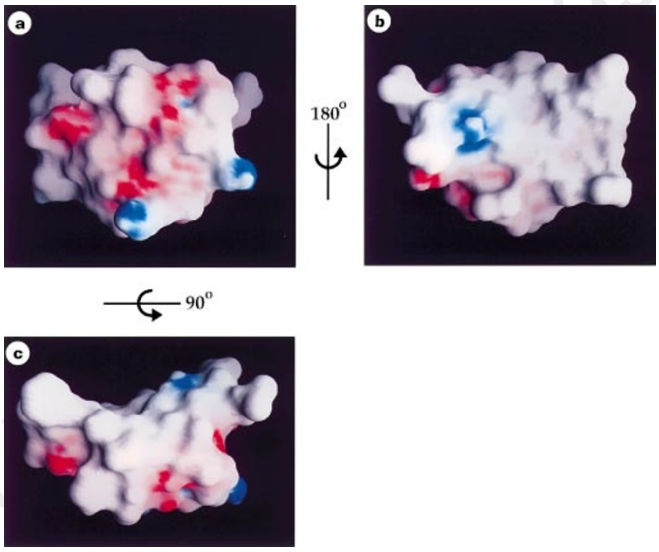

Figure 4 Surface contours and charge density. **a**, This orientation of LR5 is similar to that in Fig. 2. The surface-exposed acidic residue on the left is Glu 9. The two regions of basic potential arise from Lys 31 and Lys 33. Glu 16 is above Lys 31 near the centre of this face. The Ca^{2+} binding site region, towards the upper right, contributes little to the surface charge. **b**, The hydrophobic face of the molecule is viewed after a rotation of 180° around the vertical. **c**, A 90° rotation around the horizontal from the orientation in **a** generates this top view of LR5. The hydrophobic concave face is towards the top of the figure. Regions of basic potential (>16k_BT/e, where k_B is Boltzmann's constant and T is the absolute temperature) are shown in blue; acidic regions (<–14k_BT/e) are in red. The figure was generated using GRASP[27].

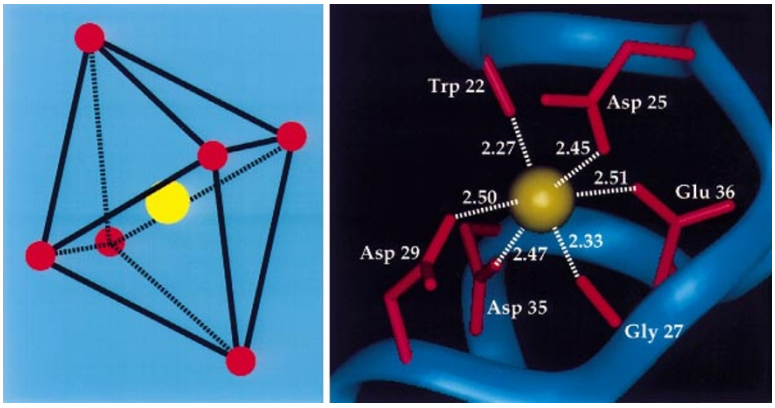

Figure 3 LR5 Ca^{2+} coordination compared to ideal octahedral coordination. Red balls indicate positions of coordinating oxygen atoms; the yellow ball shows the position of the Ca^{2+} ion. Lines are drawn between coordinating atoms in the schematic diagram to show how a six-ligand coordination geometry generates an octahedron. Side chains involved in Ca^{2+} coordination are superimposed on a ribbon trace of the LR5 backbone. Distances in Ångströms between coordinating oxygens and the Ca^{2+} ion are indicated. The figure was generated using Ribbons[25]. The distances measured in the refined LR5 structure between the side-chain carboxylate groups and the Ca^{2+} ion are 2.48 ± 0.03 Å. The distances between the backbone carbonyl oxygens and the Ca^{2+} ion are 2.30 ± 0.03 Å. The r.m.s.d. for the six oxygen atoms coordinating the Ca^{2+} ion, as compared to an ideal octahedron with centre-to-vertex distances of 2.40 Å, is 0.21 Å, comparable to deviations seen for Ca^{2+} sites in other proteins[26]. The octahedral coordination cages the Ca^{2+} ion such that the folded domain structure is likely to inhibit exchange of the Ca^{2+} ion with solvent, consistent with the ~70 nM Ca^{2+}-binding affinity of the domain[9].

of apolipoproteins[2,12–14]. However, our results show that the residues in the conserved acidic motif of LR5 are buried to participate in Ca^{2+} coordination, instead of being exposed on the surface of the domain (Fig. 4a). Although lipoprotein uptake by members of the LDLR family may involve an electrostatic component, perhaps through association of the lipoproteins with cell-surface proteoglycans[15], the LR5 structure demonstrates that the primary role for the conserved acidic residues in LDL-A modules is structural. It has been noted previously that only lipid-associated apolipoproteins bind with high affinity to the LDLR[16]; the LR5 structure suggests an alternative model for LDLR function in which a hydrophobic concave face (Fig. 4b, c) provides a lipoprotein-binding surface. □

Methods

Protein crystallization. Protein was expressed, purified and refolded as described[9]. Protein stock solution contained 10 mg ml^{-1} LR5 and 5 mM $CaCl_2$. Crystals of LR5 were grown by the hanging-drop method at 15 °C by equilibrating over a reservoir containing 2.1 M ammonium sulphate, adjusted to pH 5.0, and 25% sucrose. These crystals belong to the rhombohedral space group R3 ($a = b = c = 32.04$ Å, $\alpha = \beta = \gamma = 112.57°$), with a single LR5 molecule in the asymmetric unit. Solvent content in the crystals is 28%.

Data collection, structure determination and structure refinement. Diffraction data were collected using a rotating anode source (Rigaku RU200) with an RAXIS IIc detector. All data were collected at −180 °C using an X-stream cryogenic crystal cooler (MSC). Crystals were mounted in a 20 μm rayon fibre loop and flash-frozen in the gaseous nitrogen stream. Reflections were indexed and integrated with the program DENZO[17] and were scaled using SCALEPACK[17]. Subsequent data manipulations were carried out using the CCP4 package[18].

Heavy-atom derivatization was accomplished by collecting crystals into reservoir solution, adding solid heavy atom to 1 mM, and soaking overnight. Heavy-atom sites were located for each derivative by Patterson methods. Heavy-atom refinement and phasing were carried out with MLPHARE (CCP4). Density modification was done on the MIR map using the program DM[19]. The 3-Å density-modified map was displayed with the program O (ref. 20) for tracing of the polypeptide chain, using the NMR structures of rLB1 and rLB2 (refs 7, 8) as a guide. Phases from the initial polyalanine model were combined with MIR phases using SIGMAA weighting[21]; side chains were readily built into the phase-combined map. A set of reflections (~7%) were held aside before refinement to monitor a free *R*-factor[22]. The model was refined against the remaining data to 1.7 Å with the program XPLOR[23]. *B*-factor refinement was done first, followed by simulated annealing. A bulk solvent correction was then applied and ordered waters were added.

Received 26 March; accepted 19 May 1997.

1. Goldstein, J. L. *et al.* Receptor-mediated endocytosis: concepts emerging from the LDL receptor system. *Annu. Rev. Cell Biol* **1,** 1–39 (1985).
2. Brown, M. S. & Goldstein, J. L. A receptor-mediated pathway for cholesterol homeostasis. *Science* **232,** 34–47 (1986).
3. Esser, V. *et al.* Mutational analysis of the ligand binding domain of the low density lipoprotein receptor. *J. Biol. Chem.* **26,** 13282–13290 (1988).
4. Russell, D. W., Brown, M. S. & Goldstein, J. L. Different combinations of cysteine-rich repeats mediate binding of low density lipoprotein receptor to two different proteins. *J. Biol. Chem.* **264,** 21782–21688 (1989).
5. Altschul, S. F. *et al.* Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410 (1990).
6. Goldstein, J. L., Hobbs, H. H. & Brown, M. S. in *The Metabolic and Molecular Bases of Inherited Disease* (eds Scriver, C. S., Beandet, A. L., Sly, W. S. & Valle, D.) 1981–2030 (McGraw Hill, New York, 1995).
7. Daly, N. L. *et al.* Three-dimensional structure of a cysteine-rich repeat from the low-density lipoprotein receptor. *Proc. Natl Acad. Sci. USA* **92,** 6334–6338 (1995).
8. Daly, N. L., Djordjevic, J. T., Kroon, P. A. & Smith, R. Three-dimensional structure of the second cysteine-rich repeat from the human low-density lipoprotein receptor. *Biochemistry* **34,** 14474–14481 (1995).
9. Blacklow, S. C. & Kim, P. S. Protein folding and calcium binding defects arising from familial hypercholesterolemia mutations of the LDL receptor. *Nature Struct. Biol.* **3,** 758–762 (1996).
10. van Driel, I. R., Goldstein, J. L., Südhof, T. C. & Brown, M. S. First cysteine-rich repeat in ligand-binding domain of low density lipoprotein receptor binds Ca^{2+} and monoclonal antibodies, but not lipoproteins. *J. Biol. Chem.* **262,** 17443–17449 (1987).
11. Abbadi, A. *et al.* Involvement of side functions in peptide structures: the Asx turn. Occurrence and conformational aspects. *J. Am. Chem. Soc.* **113,** 2729–2735 (1991).
12. Südhof, T. C., Goldstein, J. L., Brown, M. S. & Russell, D. W. The LDL receptor gene: a mosaic of exons shared with different proteins. *Science* **228,** 815–822 (1985).
13. Mahley, R. W. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science* **240,** 622–630 (1988).
14. Wilson, C. *et al.* Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E. *Science* **252,** 1817–1822 (1991).
15. Ji, Z.-S., Fazio, S., Lee, Y.-L. & Mahley, R. W. Secretion-capture role for apolipoprotein E in remnant lipoprotein metabolism involving cell surface heparan sulfate proteoglycans. *J. Biol. Chem.* **269,** 2764–2772 (1994).
16. Innerarity, T. L., Pitas, R. E. & Mahley, R. W. Binding of arginine-rich (E) apoprotein after recombination with phospholipid vesicles to the low density lipoprotein receptors of fibroblasts. *J. Biol. Chem.* **254,** 4186–4190 (1979).
17. Otwinowski, Z. & Minor, W. *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. & Bailey, S.) 556–562 (SERC Daresbury Laboratory, Warrington, UK, 1993).
18. CCP4. The SERC (UK) Collaborative Computing Project No. 4 *A Suite of Programs for Protein Crystallography* (SERC Daresbury Laboratory, Warrington, UK, 1979).
19. Cowtan, K. D. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography **31,** 34–38 (1994).
20. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
21. Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42,** 140–149 (1986).
22. Brünger, A. T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355,** 472–475 (1992).
23. Brünger, A. T. *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
24. Yamamoto, T. *et al.* The human LDL receptor: a cysteine-rich protein with multiple *Alu* sequences in its mRNA. *Cell* **39,** 27–38 (1984).
25. Carson, M. & Bugg, C. E. Algorithm for ribbon models of proteins. *J. Mol. Graph.* **4,** 121–122 (1986).
26. McPhalen, C. A., Strynadka, N. C. J. & James, M. N. G. Calcium binding sites in proteins: a structural perspective. *Adv. Prot. Chem.* **42,** 77–144 (1991).
27. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11,** 281–296 (1991).

Acknowledgements. We thank S. J. Gamblin for technical advice and discussions. J.M.B. is a Whitehead Fellow and acknowledges support from the W. M. Keck Foundation. This research utilized the W. M. Keck Foundation X-ray Crystallography Facility at the Whitehead Institute, and was supported by an NIH grant (to R. D. Rosenberg) from the Program of Excellence in Molecular Biology.

Correspondence and requests for materials should be addressed to P.S.K. Coordinates have been deposited in Brookhaven Protein Data Bank under accession code 1AJJ.

Crystal structure of a small G protein in complex with the GTPase-activating protein rhoGAP

Katrin Rittinger*†, Philip A. Walker*†, John F. Eccleston*, Kurshid Nurmahomed*, Darerca Owen‡, Ernest Laue‡, Steven J. Gamblin* & Stephen J. Smerdon*

* *National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK*

‡ *Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK*

† *These authors contributed equally to this work.*

Small G proteins transduce signals from plasma-membrane receptors to control a wide range of cellular functions[1,2]. These proteins are clustered into distinct families but all act as molecular switches, active in their GTP-bound form but inactive when GDP-bound. The Rho family of G proteins, which includes Cdc42Hs, activate effectors involved in the regulation of cytoskeleton formation, cell proliferation and the JNK signalling pathway[3–9]. G proteins generally have a low intrinsic GTPase hydrolytic activity but there are family-specific groups of GTPase-activating proteins (GAPs) that enhance the rate of GTP hydrolysis by up to 10^5 times[10,11]. We report here the crystal structure of Cdc42Hs, with the non-hydrolysable GTP analogue GMPPNP, in complex with the GAP domain of p50rhoGAP at 2.7 Å resolution. In the complex Cdc42Hs interacts, mainly through its switch I and II regions, with a shallow pocket on rhoGAP which is lined with conserved residues. Arg 85 of rhoGAP interacts with the P-loop of Cdc42Hs, but from biochemical data and by analogy with the G-protein subunit $G_{i\alpha1}$ (ref. 12), we propose that it adopts a different conformation during the catalytic cycle which enables it to stabilize the transition state of the GTP-hydrolysis reaction.