Structure of the LDL Receptor Extracellular Domain at Endosomal pH

Gabby Rudenko,[1] Lisa Henry,[1,3] Keith Henderson,[4]
Konstantin Ichtchenko,[2]* Michael S. Brown,[2]
Joseph L. Goldstein,[2] Johann Deisenhofer[1,3]†

The low-density lipoprotein receptor mediates cholesterol homeostasis through endocytosis of lipoproteins. It discharges its ligand in the endosome at pH < 6. In the crystal structure at pH = 5.3, the ligand-binding domain (modules R2 to R7) folds back as an arc over the epidermal growth factor precursor homology domain (the modules A, B, β propeller, and C). The modules R4 and R5, which are critical for lipoprotein binding, associate with the β propeller via their calcium-binding loop. We propose a mechanism for lipoprotein release in the endosome whereby the β propeller functions as an alternate substrate for the ligand-binding domain, binding in a calcium-dependent way and promoting lipoprotein release.

The low-density lipoprotein receptor (LDL-R) regulates cholesterol homeostasis in mammalian cells. LDL-R removes cholesterol-carrying lipoproteins from plasma circulation in a process known as receptor-mediated endocytosis (1). Ligands bound extracellularly by LDL-R at neutral pH are internalized and then released in the endosomes (pH < 6), leading to their subsequent lysosomal degradation. The receptor then recycles to the cell surface. Mutations in the LDL-R gene cause familial hypercholesterolemia (FH), one of the most common simply inherited genetic diseases (2). FH heterozygotes exhibit a reduced rate of receptor-mediated removal of plasma LDL by the liver, ultimately leading to early onset coronary heart disease and atherosclerosis. More than 920 mutations in LDL-R are known, some of which have been functionally characterized (2, 3).

The extracellular domain of LDL-R is composed of a "ligand-binding domain" (with cysteine-rich repeats R1 to R7) and an "epidermal growth factor (EGF) precursor homology domain" (with the EGF-like repeats A, B, and C, as well as a β propeller between B and C) (4, 5). LDL-R binds LDL via the single protein in LDL, the 550-kD apolipoprotein B (apoB) (6); deleting R3, R4, R5, R6, or R7 reduces LDL binding to <20% of that of the wild-type LDL-R (7). LDL-R also binds to very low density lipoprotein (VLDL), β-VLDL, interme-

diate density lipoprotein (IDL), and chylomicron remnants via the 33-kD apolipoprotein E (apoE) (8, 9); disrupting R5 decreases β-VLDL binding to 30 to 50% of that of the wild-type receptor, whereas disrupting R4 or R6 reduces binding only slightly (7). At neutral pH, negative charges on repeats R1 to R7 are thought to interact with positive charges on apoB and apoE. Indeed, LDL binding to LDL-R can be disrupted competitively with polycations or permanently by selective chemical modification of positively charged residues on apoE or apoB (1, 10, 11). Further, the LDL-R cysteine-rich repeats contain clusters of acidic residues implicated not only in ligand binding but also in calcium binding, as shown at pH = 5.0 crystallographically by Fass et al. (12) and at pH = 6.7 with the use of nuclear magnetic resonance (NMR) imaging (13). The calcium ions enable the repeats to bind to the ligand (14, 15).

Dissociation of ligands is crucial for receptor recycling and hence proper receptor function; mutations in LDL-R that impair ligand release produce FH (2). Deletion mutagenesis studies in LDL-R and the related VLDL-R have indicated that, although the ligand-binding domain is sufficient for binding lipoprotein particles, the receptor requires the EGF precursor homology domain for ligand release (16–18).

The structural basis for LDL-R's ability to recognize a diverse group of lipoprotein particles, all varying in size, and release them at acidic pH is unknown. High-resolution crystal structures of modules R5 (12) and β propeller–C (5) are known, and solution NMR structures are known for single and tandem repeats, including R1, R2, R5, R6, A, and B (13, 19–26). These analyses have been carried out at a wide range of pH values (from 3.9 to 7.5), with most of the NMR studies done at acidic pH. Electron microscopic studies have imaged a variety of shapes (27). Here we present the crystal structure of the extracellular portion of LDL-R at

pH = 5.3 and at 3.7 Å resolution; this structure should represent the conformation of LDL-R adopted in endosomes.

Structure determination. The extracellular domain of human LDL-R (residues 1 to 699) was crystallized at pH = 5.3, with the symmetry of space group P3$_1$21 (28). Soaking crystals in sodium 12-tungstophosphate (Na$_3$PW$_{12}$O$_{40}$) improved their diffractive quality and incorporated large anomalous scatterers. The asymmetric unit contains a single protein molecule and two tungsten clusters as well as half of a tungsten cluster on a crystallographic twofold axis. Data collection, structure determination, and model statistics are given in Table 1 and in (29). The combination of good experimental phases and well-determined structures for most modules enabled us to build a reliable model at 3.7 Å that indicates the relative orientation of the modules; however, a discussion of atomic-level interactions (exact side-chain conformations) at this resolution is not possible.

Monomer description. There is clear electron density in the crystal structure for the modules R2, R3, R4, R5, R6, R7, A, B, β propeller, and C (Fig. 1A); module R1 is present but not seen because of disorder in the crystal. The modules of the ligand-binding domain (R2 to R7) form an arc that folds back onto the EGF precursor homology domain, a linear array of modules A, B, the β propeller, and C (Fig. 1, B and C). The arc from R2 to R7 (Phe^{65} to Met^{277}) spans 140 Å, and the array of A, B, β-propeller, and C (Met^{277} to Ser^{442}) spans 105 Å. The height of the molecule is 90 Å (Fig. 1B, Gln^{170} to Thr^{649}). The density for two large carbohydrate structures and the linkers between the modules are well resolved, with the exception of the linker between R4 and R5. The modules have the same fold as observed in previously determined (related) fragments. Each cysteine-rich repeat contains two loops connected by three disulfide bonds. The second loop carries conserved acidic residues and forms a calcium-binding site with the consensus sequence X$_n$ (main chain carbonyl)-X-X-**Asp/Asn**$_{n+3}$ (side chain)-X-**X**$_{n+5}$ (main chain carbonyl)-X-**Asp/Asn/Glu**$_{n+7}$ (side chain)-X-X-X-X-X-**Asp**$_{n+13}$ (carboxylate)-**Asp/Glu**$_{n+14}$ (carboxylate), with the functional groups given in parentheses, X indicating any residue, and n indicating the residue number. The A, B, and C modules also contain three disulfide bonds each. Consistent with the structural studies on fragments, the modules R2, R3, R4, R5, R6, R7, A, and B appear to contain calcium ions in their calcium-binding sites. These ions were identified as features in the electron density that far exceed the average protein density. The majority of the extracellular domain of LDL-R is thus well ordered in the crystal structure.

In the crystal, each monomer forms major contacts with five neighboring symmetry-related molecules. Although the relative orien-

[1]Department of Biochemistry, [2]Department of Molecular Genetics, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard Y4-206, Dallas, TX 75390, USA. [3]Howard Hughes Medical Institute, [4]Berkeley Center for Structural Biology, Lawrence Berkeley Laboratory, MS 6R2100, 1 Cyclotron Road, Berkeley, CA 94720, USA.

*Present address: New York University School of Medicine, 550 First Avenue, New York, NY 10016, USA.
†To whom correspondence should be addressed. E-mail: Johann.Deisenhofer@UTSouthwestern.edu

tation of some of the LDL-R modules is determined by their intermolecular interactions, other modules are positioned by their intramolecular interactions, and these interactions should also occur in solution.

The modules in the ligand-binding domain (R2 to R7), apart from their linkers, do not interact with each other directly. The linkers are short, four or five residues, except for linker R4-R5, which is 12 residues. R2 contacts module B of a symmetry-related molecule, whereas R3 makes direct contact with the β propeller of another symmetry mate. Tungsten clusters (10 Å in diameter) stabilize these interactions, one packing between R2 and R3 and another packing between R3, β propeller, and a neighboring β propeller (Fig. 2). R4 and R5 bind side by side to one end of the β-propeller "barrel" via extensive interactions, and the protein moieties do not contact each other, symmetry-related molecules, or tungsten clusters. R6 makes extensive crystal contacts to two neighboring molecules (β propeller and C module of one molecule and R7 of another). Lastly, R7 is held in place by module A and a symmetry-related molecule (side of β-propeller barrel and R6 module). Thus, in the ligand-binding domain, whereas R2, R3, and R6 are fixed by crystal contacts and R7 makes intramolecular as well as crystal contacts, R4 and R5 are held in place solely by β propeller.

The modules of the EGF precursor homology domain interact with each other in addition to contacting the ligand-binding domain. The junctions between R7, A, B, the β-propeller fold, and C direct their relative orientations. At the kink formed by R7 and A, where the ligand-binding domain folds back on the EGF precursor homology domain, hydrophobic areas of side chains pack together to form a core [disulfide Cys^{275}-Cys^{292} (of R7), Val^{274} (R7), Lys^{312} (A), and Ile^{313} (A)]. At the next junction, the carboxyl-terminal end of A interacts tightly with the amino-terminal part of B. The linker residues are sandwiched on one side by a loop from A and on the other side by a loop from B. In the middle of the linker, Asp^{333} (of A) ligands to the calcium ion in B, and Phe^{323} (A) and Glu^{351} (B) pack on either side. Module A does not make contact with symmetry-related molecules and is held in place by R7 and B. The B module is packed against the β propeller but also contacts the R2 of a neighboring molecule. Although the individual A and B modules are similar to those observed by NMR spectroscopy on A-B fragments (25), sharing 2.6 and 2.5 Å root mean square deviation (rmsd) for 39 and 41 C^{α} atoms, respectively, their relationship in the extracellular domain is different. A and B still form an extended rod, but B has swiveled about 40° with respect to A in a rigid body motion (roughly around the rod axis). Three loops from the carboxyl-terminal half of module B pack against the β propeller (Gln^{338} to Gln^{345}, Glu^{359} to Phe^{362}, and Ala^{370} to Gly^{375}) to form an extensive interface. The polypeptide chain reverses direction between B and β propeller with a sharp kink at Gly^{375}, packing hydrophobic residues together [Phe^{362} (of B) and Leu^{637} and Phe^{593} (β propeller)]. A mutation of Gly^{375} to Ser^{375} is found in FH patients (3), suggesting that nonglycine residues are incompatible with the main chain conformation. A number of interacting charged residues are buried at the interface between B and β propeller: Asp^{339} (of B) or Asp^{341} (B) with Arg^{553}, Asp^{548}, Glu^{594}, and Arg^{612} (β propeller). The mutations of Asp^{341} to Ala^{341}, Gln^{345} to Arg^{345}, Arg^{553} to Cys^{553} (R553C), Glu^{594} to Lys^{594}, and Arg^{612} to Cys^{612} (R612C) are found in FH patients (3); all five mutations would change the balance and/or location of charges, likely disrupting the B–β propeller interface, whereas R553C and R612C might also affect folding and proper disulfide bond formation. The presence of additional modules (R1 to R7 and A) in our structure and/or the acidic pH seems to promote the interface between B and β propeller, because in the crystal structure of the β propeller–C fragment at neutral pH the B module is present but disordered (5). The C module packs against the β propeller as was seen at neutral pH in the crystal structure of the fragment (5). The modules in the EGF precursor homology domain observed in our crystal structure influence each other by imposing relative orientations through their junctions; they influence the ligand-binding domain as well.

The electron density for linker R4-R5 is ambiguous in our crystal structure due to flexibility or static disorder. The R4-R5 linkers from two different monomers pack on either side of a two-fold crystallographic axis, so that module R4 could conceivably connect to module R5* instead of module R5 and coexist in the crystal (Fig. 2). Linker length, electron density, and the monomeric state observed in gel filtration studies at pH 5 make a domain-swapped scenario unlikely. In any case, chemical interactions between R4, R5, and the β-propeller fold would not change, and our observations would remain the same.

Our crystal structure reflects three important aspects of the LDL-R molecule: (i) The loops in

Fig. 1. (**A**) The experimental electron density calculated with phases from the MAD experiment after density modification. (To portray the density around a single monomer, especially between the modules, in a complicated crystal packing, density farther than 1.8 Å from the protein atomic coordinates has been omitted.) The quality of the map was further improved by incorporating model phases and by sharpening observed structure factor amplitudes (not shown). (**B**) C^{α} trace of LDL-R monomer. Modules are colored according to their boundaries with the ligand-binding domain containing R2 (residues 44 to 85), R3 (85 to 124), R4 (124 to 170), R5 (170 to 212), R6 (212 to 254), R7 (254 to 294), and the EGF precursor homology domain containing A (294 to 332), B (332 to 377), β propeller (377 to 643), and C (643 to 693). Regions of poor backbone connectivity are dashed. Calcium ions are indicated as red spheres, and disulfide bonds and carbohydrates on Asn^{135} and Asn^{251} are shown in gray as ball and stick (sulphur atoms, yellow; oxygen, red; nitrogen, blue; and carbon, gray). (**C**) The view is rotated 90° from that of (B).
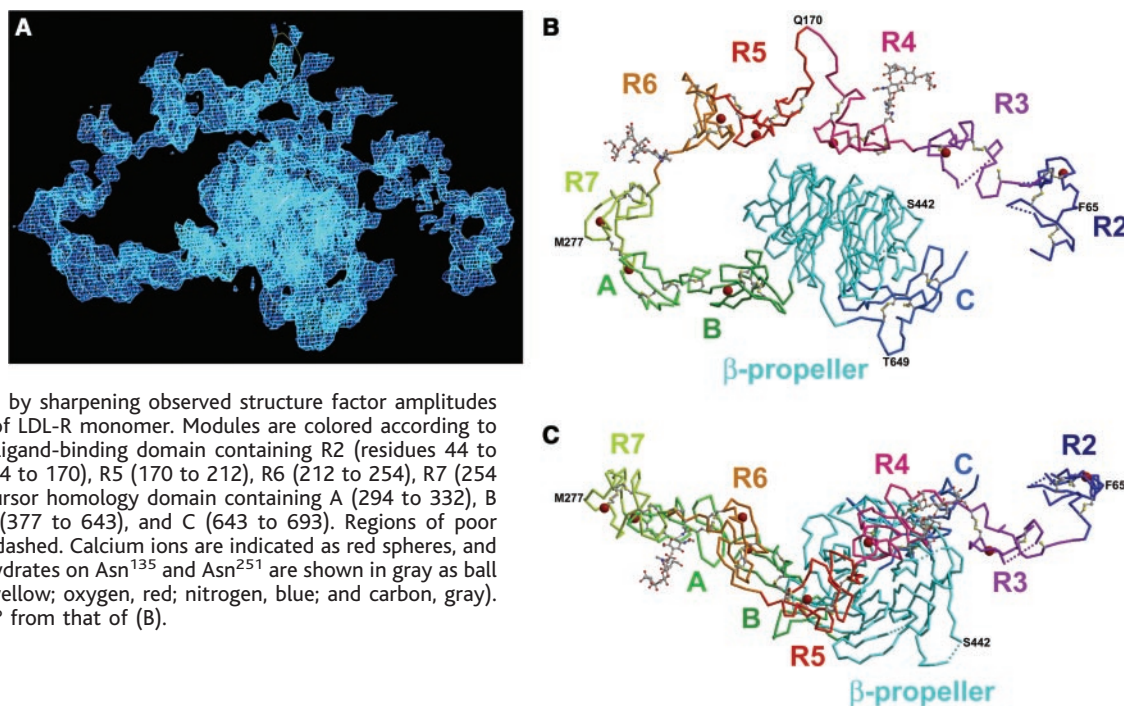
the cysteine-rich repeats, with their lack of secondary structure elements, are pliable. They mold in different ways to form crystal contacts with apparent ease, as no helical or β-sheet interactions need to be broken. In the NMR studies, this flexibility has manifested itself in ensembles of widely divergent conformations. (ii) The length and flexibility of the linkers between the cysteine-rich repeats enable them to adjust with respect to each other under the influence of other modules or crystal contacts. R5 and R6 fold toward each other and R3 and R4

extend away, whereas R4 and R5 align side by side. (iii) At endosomal pH, some modules are constrained with respect to each other (R4, R5, R7, A, B, β propeller, and C), whereas others apparently are not.

Interface of R4 and R5 with β propeller. The modules R4 (residues Cys[127] to Cys[163]) and R5 (Cys[176] to Cys[210]) bind β propeller (Ile[377] to Gly[642]) via extensive interactions (Fig. 3). The calcium-binding loops of both R4 and R5 dock side by side, one on the hub and one to the side of the six-bladed β-pro-

peller barrel, placing the calcium ions circa (ca.) 21 Å apart. In contrast, the calcium-binding sites in the other repeats point in different directions, exposed to the solvent (R6 and R7) or (partially) involved in crystal contacts (R2 and R3). Although the precise location of the amino acid side chains cannot be determined given the resolution of the diffraction data, the following nevertheless appears clear: In R4, the calcium-binding loop provides the chelating residues Trp[144], Asp[147], Asp[149], and Asp[151]— in addition to Ile[140], Pro[141], and Pro[150]— to the interface. Hydrophobic contacts are made between R4 (Ala[130], Ile[140], Pro[141], Trp[144], and Asp[151]) and aliphatic areas on β propeller (Trp[515], Thr[517], Gln[540], Trp[541], Lys[560], and His[586]). Salt bridges are formed as well (Asp[147] to Lys[560] and Asp[149] to His[562] and His[586]). In R5, the calcium-binding ligands Trp[193], Asp[196], Gly[198], and Asp[200]— in addition to His[190], Ser[191], and Lys[202] from the same loop— also bind β propeller. Hydrophobic interactions are made (Trp[193] between Glu[581] and Lys[582]), and salt bridges are formed (His[190] to Glu[581], Asp[196] and Asp[200] to Lys[582], and likely Lys[202] to Glu[623]). A cluster of histidines (His[190], His[562], and His[586]) is found at the three-way junction between R4, R5, and β propeller. All together, R4 and R5 provide β propeller with 14 major residues to form an interface mediated by hydrophobic and charged interactions.

The relative dispositions of R4, R5, and β propeller appear highly specific, and in particular the R4–β propeller interface is highly conserved. Eleven out of 14 residues from R4 and
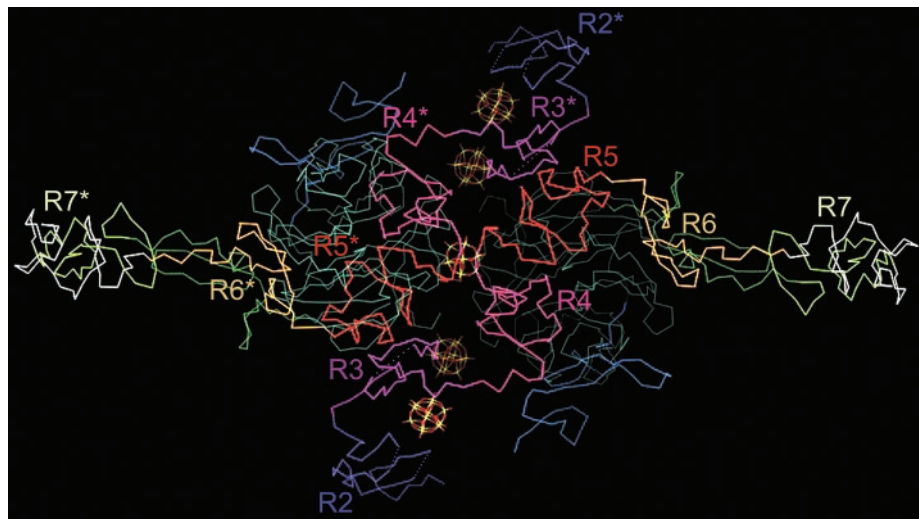

Fig. 2. Two monomers stabilized by tungsten clusters; the ligand-binding domains have been labeled R2 to R7 and R2* to R7*, respectively. The linkers between R4 and R5 and R4* and R5* meet at a crystallographic twofold axis. At a distance below, a cluster (center) sits on the twofold as well.

Table 1. Statistics from the crystallographic structure determination (space group P3$_1$21). HER, high energy remote; LER, low energy remote; and Infl, inflection. The number of crystals used per data set is given in parentheses after the unit cell dimensions. Outer-shell statistics are given in parentheses for data resolution, data completeness, $R_{merge(I+, I-)}$, $R_{merge(I)}$, and $<<I>/\sigma(<I>)>$ {$\sigma(I)$ is the corrected standard deviation [see SCALA/CCP4 (43)]}. Entries marked by a dash are not relevant. Completeness in outer shells is low because of anisotropic diffraction and subsequent anisotropic integration. $R_{merge(I+, I-)}$ is the multiplicity weighted R_{merge} keeping $I+$ and $I-$ separate [see SCALA/CCP4 (43)]. Bijvoet and symmetry-related pairs were kept together for scaling but separated for merging in peak, inflection, and HER, but not LER,

data. $R_{merge(I)}$ is the multiplicity weighted R_{merge} with respect to I_{mean} (i.e., $I+$, $I-$, and their symmetry-related mates). The traditionally reported I/σ (where σ is the rms scatter of observations) are as follows (outer-shell statistics are in parentheses): peak data, $I/\sigma = 3.2$ (2.5); inflection data, $I/\sigma = 4.1$ (1.9); HER, $I/\sigma = 4.8$ (1.6); and LER, $I/\sigma = 2.6$ (2.0). The data show a more or less linear fall-off and are weak. Wilson B-factor values calculated by TRUNCATE/CCP4 (43) from lowest to highest resolution shell exemplify this: peak, 133 Å²; inflection, 122 Å²; HER, 123 Å²; LER, 125 Å². R_{ano} is given by $<||F+| - |F-||>/<\frac{1}{2}(|F+| + |F-|)>$ as calculated by CNS (46). R_{disp} is given with respect to inflection, HER, or LER as $<| |F(\lambda2)| - |F(\lambda1)| |>/<|(\overline{F(\lambda1)} + |F(\lambda2)|/2>$ as calculated by CNS (46). R_{cullis} and phasing power are as given by SHARP (45).

Data set	λ(Å)	Unit cell (Å)	Resolution (Å)	Total/unique reflections	Completeness (%)	$R_{merge(I+, I-)}$ (%)	$R_{merge(I)}$ (%)	$<<I>/\sigma(<I>)>$
Peak	1.21363	184.80 184.80 84.96 (3)	39.2–3.8 (3.93–3.80)	101571/14057	84.6 (20.7)	11.4 (39.6)	19.3 (42.5)	10.1 (1.6)
Inflection	1.21423	184.80 184.80 84.88 (1)	49.4–4.0 (4.14–4.00)	43739/13002	91.4 (51.2)	12.0 (50.1)	17.5 (49.0)	7.7 (2.4)
HER	1.10696	184.21 184.21 84.72 (1)	45.2–3.8 (3.93–3.80)	47496/14789	89.6 (32.0)	9.5 (59.6)	11.4 (55.6)	8.1 (2.4)
LER	1.23985	185.29 185.29 85.19 (8)	33.7–3.7 (3.83–3.70)	64801/17303	95.6 (64.5)	–	9.0 (44.9)*	9.8 (2.7)

MAD phasing statistics

	Resolution (Å)	R_{ano}	R_{disp} (Infl HER LER)			R_{cullis} (acentric/centric) Δiso	Δano	Phasing power (acentric/centric) Δiso	Δano
Peak	50–3.8	0.166	0.12	0.12	0.15	0.75/0.82	0.53/–	1.03/0.75	5.24/–
Inflection	50–4.0	0.142		0.14	0.17	0.80/0.78	0.69/–	2.39/1.42	3.67/–
HER	50–3.8	0.106			0.12	0.63/0.55	0.77/–	0.60/1.28	2.78/–
LER	50–3.7	–				–/–	0.71/–	0.0/0.0	1.81/–

Figure of merit to 3.7 Å (acentrics/centrics): 0.64/0.53 calculated for 31 W atoms per asymmetric unit, including 2 W atoms on special positions.
0.34/0.25 (3.95 to 3.7 Å)

*The nonmultiplicity weighted (traditional) R_{merge} for LER is 7.7% (38.8% in the outer shell).

R5 contacting β propeller are identical or highly conserved in nine LDL-R sequences and six VLDL-R sequences (the equivalent modules R5 and R6) (Fig. 4A, top). Of the 10 counterpart residues on β propeller, 6 are identical or highly conserved, and at two other sites adjacent residues in the sequence could provide similar chemical interactions (not shown). Structural features such as the three salt bridges between R4 and β propeller and the cluster of histidines are conserved in the 15 sequences, as well as three out of four salt bridges between R5 and the β-propeller fold in many species, suggesting

functional relevance. In fact, the region obscured by R4 and R5 on β propeller belongs to the largest area of contiguous sequence conservation between LDL-R and VLDL-R on the molecular surface (Fig. 4B).

To confirm that the ligand-binding domain binds to the β-propeller fold at endosomal pH in solution, we cleaved the protein between R4 and R5 with α-chymotrypsin at pH = 6 and pH = 8 and performed size exclusion studies. At pH = 6, the 33-kD fragment R1-R4 remains associated with the rest of the molecule (Fig. 5; lanes 1, 2, and 3). At pH = 8, the fragment R1-R4

dissociates (Fig. 5, lanes 6 and 7). When the cleavage reaction is performed at pH = 5.5 and the pH subsequently raised to 8, the fragment R1-R4 dissociates, whereas cleaving LDL-R in two at pH = 8 followed by a shift to acidic pH results in R1-R4 being retained with the rest of the molecule (30).

Implications of the LDL-R crystal structure—a mechanism for ligand release at acidic pH. The compact shape of LDL-R seen at pH = 5.3 confirms the protein properties observed during our quest for crystals. Gel filtration studies at pH = 5 indicated a hydrodynamic radius corresponding to a monomer with an apparent mass of 95 kD. In contrast, at pH = 8 and in the presence of either $CaCl_2$ or EDTA, gel filtration and dynamic light-scattering studies revealed a hydrodynamic radius too small for a dimer unless completely globular and more consistent with an elongated monomer (apparent mass of 160 kD). Studies with other techniques have been hampered by the low solubility of LDL-R at acidic pH. The association of R1-R4 with the rest of the molecule was suspected early, because proteolysis of the linker R4-R5 took place on occasion during crystal

Fig. 3. Interface formed between R4, R5, and β propeller. Residues discussed in the text are shown and labeled except for Ser[191], which was omitted for clarity. R4, magenta trace; R5, red trace; β propeller, cyan trace; and calcium ions, maroon spheres.
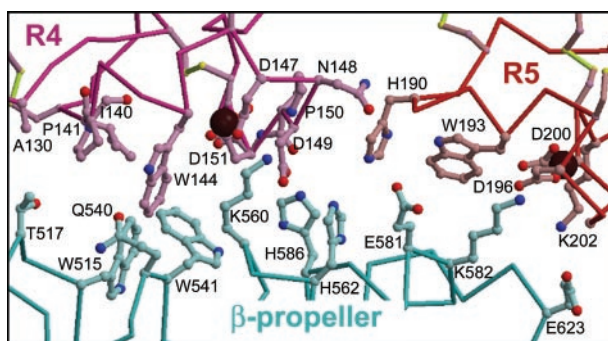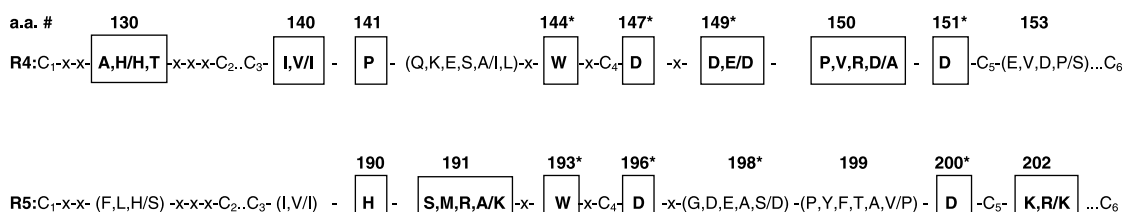

Fig. 4. (**A**) Fifteen sequences for LDL-R [human, rabbit, rat, mouse, pig, frog (two genes), hamster, shark] and VLDL-R (human, mouse, rabbit, rat, cow, chick) were aligned with CLUSTAL W (48). The different LDL-R sequences share 66 to 85% identity to human LDL-R. The VLDL-R sequences share ca. 53% identity to human LDL-R. (Top) Sequence homology for R4 and R5, respectively, with residues given as (LDL-R/VLDL-R) (52). Residues playing a major role in forming the interface with β propeller are boxed. Calcium-binding residues are indicated with asterisks; a.a., amino acid. (Bottom) Residues in R1, R2, R3, R6, and R7, respectively, that are found at an equivalent position to the boxed residues in R4 and R5, given as (LDL-R/VLDL-R). (**B**) Sequence conservation between LDL-R and VLDL-R mapped onto the molecular surface of the human LDL-R front side (I) and rotated 180° to reveal the backside (II). The ligand-binding domain and EGF precursor homology domain have been separated at the junction of R7 and A (scissors) and rotated apart (cyan arrows) to reveal the surface buried at the interface of R4, R5, and β propeller (white arrows). Regions of sequence identity, green; semiconserved regions, yellow; and nonconserved regions, magenta. Carbohydrate surfaces are shown in white. The following groups were considered semiconserved: (I, V, L, and M), (F, Y, and W), (K, R, and H), (D, E, N, and Q), (G, A, S, and T), (C), and (P) (52). Calcium-binding sites are indicated with red arrowheads for R2 to R7.
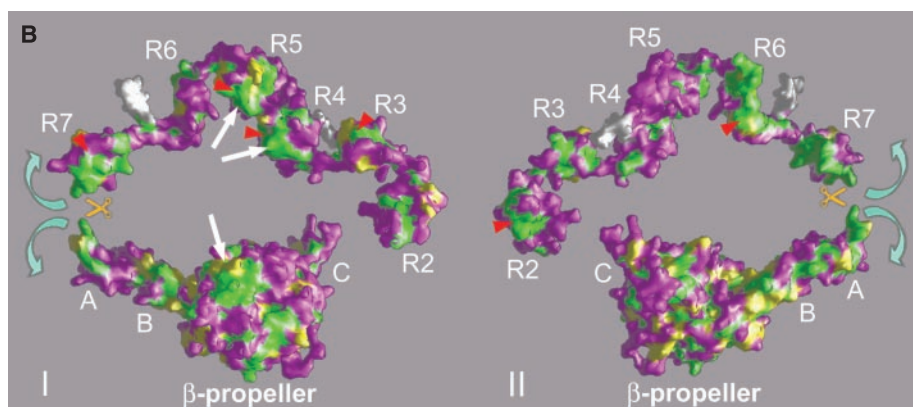
growth (confirmed by amino-terminal sequencing), yet both fragments remained in the crystal.

Our structure suggests that, at endosomal pH, β propeller might act as an alternative substrate for the ligand-binding domain. Biochemical and structural observations support this proposal. Module R5 is essential for LDL and β-VLDL binding, and R4 is important for LDL binding as well (7). In the crystal structure at acidic pH, R4 and R5 bind independently of each other to β propeller. Deleting the A and B modules creates a receptor that binds ligand but does not recycle, indicating that ligand release is not possible (31). Removing A and B would hinder R4 and R5 from reaching β propeller, and other modules are apparently not functionally equivalent (Fig. 4A, bottom). Mutant LDL-R and VLDL-R lacking the complete EGF precursor homology domain (with its β propeller) still bind ligand at the cell surface (though with reduced affinity for LDL) and are internalized but fail to recycle (16–18). Although the heparan sulfate analog suramin (binding LDL and β-VLDL) successfully dissociates lipoproteins at the cell surface in vitro from LDL-R and the mutant LDL-R lacking the EGF precursor homology domain, exposure to acidic pH only releases lipoproteins from the ligand-binding domain when the EGF precursor homology domain is present (1, 16). Many of the residues that bind the calcium ion in R4 and R5 also form the interface with β propeller, and these residues (or ones close by) likely bind to apoB and apoE as well. For example, two of these, Trp^{144} and Trp^{193}, donate only a main-chain carbonyl to calcium coordination; nevertheless their side chains are strictly conserved in 15 sequences. R6 appears different in this respect; placing a Trp at the equivalent site appears detrimental

(Fig. 4A, bottom), because the mutation from Arg^{232} to Trp^{232} in R6 is found in FH patients (3). The interface of R4, R5, and β propeller displays structural features that are well suited to support reversible, pH-regulated binding. Tryptophan residues can generate high binding affinity, as shown for the human growth hormone receptor (32). In the pH range from 5 to 7.5, the formation and strength of the interface salt bridges could change as functions of the pH and the pK_a (where K_a is the acid dissociation constant) of the aspartates, glutamates, and histidines involved. At acidic pH, histidine residues support a (partial) positive charge, explaining how a histidine cluster could stabilize buried negatively charged residues nearby. Mutations of two of the three histidines at the interface have been found in FH patients, His^{190} to Tyr^{190} and His^{562} to Tyr^{562} (3), again suggesting functional importance. Histidines have been implicated in other acid-regulated mechanisms (33, 34). For LDL-R, NMR studies of modules at acidic and more neutral pH have not identified any side chains with pH-specific conformations. Thus, it is possible that, in the endosome, the EGF precursor homology domain of LDL-R promotes ligand release by presenting the ligand-binding domain with a better substrate, the β propeller, in an intramolecular displacement reaction.

LDL-R exhibits a wide range of binding affinities to different lipoprotein particles, and the crystal structure even at acidic pH suggests a molecule accommodating ligand variation. The apoE and apoB epitopes vary with lipoprotein particle size (35, 36). The receptor-binding region of apoE forms a single bent helix that spans 70 Å in the presence of the lipid-mimicking trifluoroethanol; R5 together with a second cys-

teine-rich repeat could easily fit along the length (37). Three antibodies to apoB, which fail to recognize apoB when LDL-R is bound, map in an arc of ca. 106° on the apoB "protein belt" encircling LDL (38). Projected on a sphere with a diameter of 275 Å, the path rendered inaccessible by LDL-R would correspond to roughly 255 Å. On the basis of our structure, LDL-R in the extended form could span about 270 Å (140 Å for R2 to R7, an estimated 25 Å if R1 plus linker is included, and 105 Å added for the EGF precursor homology domain). Inasmuch as the lipoprotein particle size and content is modulated by diet, disease, and pharmacopoeia, in turn affecting its plasma clearance (1, 9, 35, 39–41, and many others), recognition of lipoproteins by LDL-R is of great clinical relevance. The extracellular domain of LDL-R appears ideally suited to recognize, bind, and release varying particles through the combination of a flexible modular ligand-binding domain and a more-or-less rigid acid-release domain.

References and Notes

1. M. S. Brown, J. L. Goldstein, Science **232**, 34 (1986).
2. J. L. Goldstein, H. H. Hobbs, M. S. Brown, in *The Metabolic & Molecular Bases of Inherited Disease*, C. R. Scriver *et al.*, Eds. (McGraw-Hill, New York, 2001), vol. II, chap. 120, pp. 2863–2913.
3. Databases for FH mutations are available online at www.ucl.ac.uk/fh and www.umd.necker.fr.
4. T. Yamamoto *et al.*, Cell **39**, 27 (1984).
5. H. Jeon *et al.*, Nature Struct. Biol. **8**, 499 (2001).
6. J. P. Segrest, M. K. Jones, N. De Loof, N. Dashti, J. Lipid Res. **42**, 1346 (2001).
7. D. W. Russell, M. S. Brown, J. L. Goldstein, J. Biol. Chem. **264**, 21682 (1989).
8. D. A. Chappell, G. L. Fry, M. A. Waknitz, L. E. Muhonen, M. W. Fladet, J. Biol. Chem. **268**, 25487 (1993).
9. A. D. Cooper, J. Lipid Res. **38**, 2173 (1997).
10. M. S. Brown, T. F. Deuel, S. K. Basu, J. L. Goldstein, J. Supramol. Struct. **8**, 223 (1978).
11. T. L. Innerarity, K. H. Weisgraber, S. C. Rall Jr., R. W. Mahley, Acta. Med. Scand. Suppl. **715**, 51 (1987).
12. D. Fass, S. Blacklow, P. S. Kim, J. M. Berger, Nature **388**, 691 (1997).
13. D. Clayton, I. M. Brereton, P. A. Kroon, R. Smith, FEBS Lett. **479**, 118 (2000).
14. W. J. Schneider, S. K. Basu, M. J. McPhaul, J. L. Goldstein, M. S. Brown, Proc. Natl. Acad. Sci. U.S.A. **76**, 5577 (1979).
15. T. Simmons, Y. M. Newhouse, K. S. Arnold, T. L. Innerarity, K. H. Weisgraber, J. Biol. Chem. **272**, 25531 (1997).
16. C. G. Davis *et al.*, Nature **326**, 760 (1987).
17. I. Mikhailenko *et al.*, J. Cell. Sci. **112**, 3269 (1999).
18. Y. Miyake, S. Tajima, T. Funahashi, A. Yamamoto, J. Biol. Chem. **264**, 16584 (1989).
19. N. L. Daly, M. J. Scanlon, J. T. Djordjevic, P. A. Kroon, R. Smith, Proc. Natl. Acad. Sci. U.S.A. **92**, 6334 (1995).
20. ———, T. L. Djordjevic, P. A. Kroon, R. Smith, Biochemistry **34**, 14474 (1995).
21. S. Bieri *et al.*, Biochemistry **37**, 10994 (1998).
22. N. Beglova, C. L. North, S. C. Blacklow, Biochemistry **40**, 2808 (2001).
23. C. L North, S. C. Blacklow, Biochemistry **39**, 2564 (1999).
24. ———, Biochemistry **38**, 3926 (1999).
25. S. Saha *et al.*, Structure **9**, 451 (2001).
26. N. D. Kurniawan, K. Aliabadizadeh, I. M. Brereton, P. A. Kroon, R. Smith, J. Mol. Biol. **311**, 341 (2001).
27. H. Jeon, G. G. Shipley, J. Biol. Chem. **275**, 30458 (2000).
28. Twenty-two mutants of the extracellular domain of LDL-R (amino acids 1 to 699) with different combinations of the five putative amino-linked glycosylation sites altered to glutamine were screened for adequate baculovirus-mediated protein overexpression in insect cells (*Trichoplusia ni*) and for crystal growth. The structure was solved with a LDL-R with the mutations of

Fig. 5. SDS–polyacrylamide gel electrophoresis of cleavage products formed by α-chymotrypsin digestion of LDL-R. a.a., amino acid. The LDL-R protein (1 mg/ml) was incubated with α-chymotrypsin (molar ratio of protease:receptor, 1:125) for 5 to 10 min at room temperature in 20 mM Tris, 100 mM NaCl, and 0.5 mM CaCl$_2$ (pH = 8) or 35 min at room temperature in 50 mM sodium acetate, 100 mM NaCl, and 0.5 mM CaCl$_2$ (pH = 6). The reactions were diluted 20- to 30-fold in buffer with 1 mM phenylmethylsulfonyl fluoride to stop proteolysis and then spun through microcon filters with a 50-kD or 100-kD molecular weight cut-off. In an unassociated form, the fragment R1-R2-R3-R4 (amino acids 1 to 167) will pass through a microcon-50 into the flow-though, but the fragment R5-R6–R7–A–B–β propeller–C (amino acids 168 to 705) and the full-length extracellular domain do not. The latter, however, do pass though a microcon-100 at pH = 8. In the pH range from 5 to 6, both the cleaved and intact protein are (reversibly) poorly soluble, and little passes through a microcon-100 (shown for the cleaved LDL-R in lane 4). Lane 1, digest at pH = 6; lane 2, flow-through from a microcon-50, pH = 6 digest; lane 3, protein retained by a microcon-50 from pH = 6 digest; lane 4, material retained by a microcon-50 from pH = 6 digest subsequently spun through a microcon-100; lane 5, markers; lane 6, digest at pH = 8, protein retained in a microcon-50; lane 7, flow-through of a microcon-50 from pH = 8 digest. The cleavage site was confirmed by amino-terminal sequencing. The gels were stained with Coomassie blue.
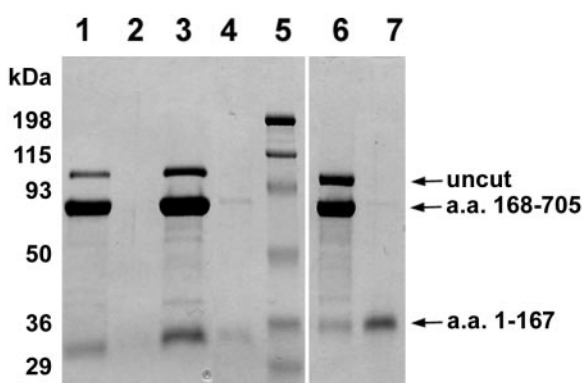
Asn[494] to Gln[494] and Asn[636] to Gln[636]. This mutant shows similar affinity for LDL as compared to the wild-type extracellular domain [with the use of a plate-binding assay similar to (15)]. Secreted carboxyl-terminally His-tagged protein was purified by Ni affinity chromatography [in 25 mM sodium phosphate (pH = 8) and 500 mM NaCl, eluted with 500 mM imidazole], Mono Q [25 mM sodium phosphate (pH = 8) and 100 mM NaCl, eluted with NaCl], and gel filtration on a Superdex-200 column [25 mM sodium phosphate (pH = 8), 150 mM NaCl] and then crystallized out of precipitate (space group P3$_1$21) with 50 mM sodium acetate (pH = 5.3), 3% 1,2 hexanediol, and 0.5 mM CaCl$_2$. The crystals contained one molecule in the asymmetric unit and 74% solvent. Micromolar amounts of sodium 12-tungstophosphate were soaked into crystals to prepare for the multiwavelength anomalous diffraction (MAD) experiment. The crystals cracked upon soaking and were radiation sensitive, so hundreds had to be screened for isomorphous pieces suitable for data collection under cryogenic conditions.

29. The structure was determined with MAD techniques. Diffraction data were collected at Advanced Light Source (ALS) (5.0.2 and 8.2.1) and Advanced Photon Source (APS) (ID-19). The data were processed with MOSFLM (42) and SCALA (43) or HKL2000 (44), followed by programs from the CCP4 suite (43) and SHARP (45) for the structure determination. Data collected at the peak, inflection, and high-energy remote wavelengths have high merging statistics as a consequence of the large anomalous signal. This was confirmed by analyzing non-Bijvoet merged data as well as by data collected at the low-energy remote wavelength. The cluster sites were derived from anomalous difference Pattersons with the use of peak data and confirmed by anomalous difference Fourier techniques with the use of molecular replacement phases obtained from the correctly positioned model pdb:1IJQ (5) done by CNS (46). The tungsten positions in the cluster were modeled on the basis of the small molecule structure (47). Experimental density–modified phases (with no model contribution or data sharpening) provided an electron density sufficient to fit the polypeptide trace for each module unambiguously. All the modules of the extracellular domain are visible except for R1 amino-

terminal sequencing of the protein from crystals shows that the first residue is Asp[4], suggesting that while R1 is present it is disordered. As a starting point for model-building, we used the fragments from PDB 1AJJ, 1D2J, 1IOU, and 1IJQ. Including the model in the phase calculation and sharpening the data by applying a resolution-dependent B factor [exp (–B/o/l^2), where d is the resolution and B = –120 Å^2, derived from the Wilson plot of the data] resolved many of the side chains. The model includes amino acids 44 to 693, except Arg[57], Val[58], Asp[75], Lys[99], Cys[100] to Ser[102], Val[388], Val[441], Gln[453], and Ala[454], for which there is poor density because of heavy atom Fourier ripples from the clusters. About 10% of the residues are currently modeled as alanines. For molecular surface calculations, all side chains were incorporated. Restrained positional refinement with a maximum likelihood target function including experimental phase information was applied with CNS to regularize the model. Crystallographic refinement is problematic because of the poor ratio of observations to parameters (with 17,303 unique reflections and more than 5000 atoms) and the lack of noncrystallographic symmetry. The R_{work} is 38.8% with the use of all reflections F > 0 between 33.1 and 3.7 Å; the R_{free} is 39.2%, calculated with a test set of 7% of the total reflections. The model has root mean square (rms) bond-length deviations of 0.03 Å and 4.9° for bond angles. The model has 583 residues in the most favored and allowed regions and 58 in the generous region, whereas six residues are currently in the unfavorable regions of the Ramachandran plot [PROCHECK (43)]. Figures were made with CLUSTAL W (48), MOL-SCRIPT (49), RASTER 3D (50), and GRASP (51).

30. G. Rudenko et al., unpublished data.
31. D. R. Van der Westhuyzen et al., Biochem. J. 277, 677 (1991).
32. T. Clackson, J. A. Wells, Science 267, 383 (1995).
33. T. Doi et al., J. Biol. Chem. 269, 25598 (1994).
34. S. Wragg, K. Drickamer, J. Biol. Chem. 274, 35400 (1999).
35. S. G. Young, Circulation 82, 1574 (1990).
36. X. Wang, R. Pease, J. Bertinato, R. W. Milne, Arterioscler. Thromb. Vasc. Biol. 20, 1301 (2000).
37. V. Raussens, C. M. Slupsky, R. O. Ryan, B. D. Sykes, J. Biol. Chem. 277, 29172 (2002).
38. J. E. Chatterton et al., J. Lipid Res. 36, 2027 (1995).

39. S. Keidar, A. C. Goldberg, K. Cook, J. Bateman, G. Schonfeld, J. Lipid Res. 30, 1131 (1989).
40. J. R. McNamara, D. M. Small, Z. Li, E. J. Schaefer, J. Lipid Res. 37, 1924 (1996).
41. Y. Toyota, T. Yamamura, Y. Miyake, A. Yamamoto, Atherosclerosis 147, 77 (1999).
42. A. G. W. Leslie, Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography (Science and Engineering Research Council Daresbury Laboratory, Warrington, UK, 1992), issue 26.
43. Collaborative Computational Project Number 4, Acta Crystallogr. D 50, 760 (1994).
44. Z. Otwinowski, W. Minor, Methods Enzymol. 276, 307 (1997).
45. E. de la Fortelle, G. Bricogne, Methods Enzymol. 276, 472 (1997).
46. A.T. Brunger et al., Acta Crystallogr. D 54, 905 (1988).
47. J.-C. P. Gabriel, R. Nagarajan, S. Natarajan, A. K. Cheetham, C. N. R. Rao, J. Solid State Chem. 129, 257 (1997).
48. J. D. Thompson, D. G. Higgins, T. J. Gibson, Nucleic Acids Res. 22, 4673 (1994).
49. P. J. Kraulis, J. Appl. Crystallogr. 24, 946 (1991).
50. E. A. Merritt, D. J. Bacon, Methods Enzymol. 277, 505 (1997).
51. A. Nicholls, K. Sharp, B. Honig, Proteins 11, 281 (1991).
52. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr; and X, any amino acid.

Formation of a One-Dimensional Array of Oxygen in a Microporous Metal-Organic Solid

Ryo Kitaura,[1] Susumu Kitagawa,[1]* Yoshiki Kubota,[2]
Tatsuo C. Kobayashi,[3] Koichi Kindo,[3] Yoshimi Mita,[4]
Akira Matsuo,[3] Michihiro Kobayashi,[4] Ho-Chol Chang,[1]
Tadashi C. Ozawa,[3] Megumi Suzuki,[4] Makoto Sakata,[5] Masaki Takata[5,6]

We report the direct observation of dioxygen molecules physisorbed in the nanochannels of a microporous copper coordination polymer by the MEM (maximum entropy method)/Rietveld method, using in situ high-resolution synchrotron x-ray powder diffraction measurements. The obtained MEM electron density revealed that van der Waals dimers of physisorbed O$_2$ locate in the middle of nanochannels and form a one-dimensional ladder structure aligned to the host channel structure. The observed O–O stretching Raman band and magnetic susceptibilities are characteristic of the confined O$_2$ molecules in one-dimensional nanochannels of CPL-1 (coordination polymer 1 with pillared layer structure).

The confinement of molecules into low-dimensional nanospace may alter their properties and reactivity, especially in the case of O$_2$ molecules, which have rich redox and magnetic properties arising from their unpaired electrons (1, 2). In spite of many ex-

perimental and theoretical investigations, the adsorption mechanism and ordering state of adsorbed O$_2$ molecules that form a specific array in nanochannels are not yet clear (3, 4). In the past decade, various novel nanochannel structures constructed with metal-organic frameworks have been produced and studied with regard to their adsorption performance with various gas molecules (5–8). If the gas molecules could be trapped in mid-channel, low-dimensional nanostructures could be designed, depending on the host channel structure.

Our strategy to elucidate the specific structure of O$_2$ molecules adsorbed in nanochannels uses a crystalline microporous copper coordination polymer that forms a uniform nanosized one-dimensional (1D) channel. We obtained accurate powder x-ray diffraction (XRD) data by synchrotron powder diffraction with a large Debye-Scherrer camera and were able to image the diffraction data by the MEM (maximum entropy method)/Rietveld method, which enabled us to determine a precise electron density map (9–12).