Model of human low-density lipoprotein and bound receptor based on CryoEM

Gang Ren, Gabby Rudenko, Steven J. Ludtke, Johann Deisenhofer, Wah Chiu, and Henry J. Pownall

Low-density lipoproteins (LDL), which are heterogeneous with respect to composition, shape, size, density, and charge (1, 2), are the major carriers of cholesterol in human plasma. LDL is removed from plasma by hepatic LDL receptors (LDLR), which maintain cholesterol homeostasis (3). LDLr variants with impaired binding to LDL cause familial hypercholesterolemia that leads to premature atherosclerotic coronary artery disease in affected patients (4). The extracellular domain of LDLr binds to LDL via apo B-100, a 4,536 amino acid polypeptide, after which the complex (LDL · LDLr) undergoes endocytosis, lysosomal degradation of LDL, and receptor recycling to the cell surface (1, 3, 4). Thus, interactions between LDL and LDLr are integral to the cholesterol homeostasis that regulates plasma LDL levels.

LDL exhibits a thermal liquid crystalline-to-isotropic transition of its cholesteryl esters (CEs) between 25 and 35 °C (5). Although several models of LDL attempt to integrate its structure and biology (5–9), a reliable three-dimensional structure of LDL or the LDL · LDLr complex has not yet been reported. Herein, we present a model of LDL and the complex, LDL · LDLr, of LDL-bound LDL receptor extracellular domain (1–699 a.a.) at extracellular pH determined by electron cryomicroscopy (cryoEM), a technique that preserves the native structure of the particles.

Result and Discussion

CryoEM micrographs of LDL embedded in vitreous ice contain spherical, ellipsoidal, and discoidal particles with internal striations (Fig. L4 and B). Since the LDL preparation is heterogeneous, our reconstruction was computed from ~8,500 particle images of LDL particles that were computationally selected from an original pool over ~48,000 particle images. The convergence of the structure from this subpopulation of particle images provides a statistically defined and robust density map displaying the most prominent and reliable structural features of LDL (Fig. 1C, and Figs. S1, 2). The reconstructed LDL subpopulation is approximately a flattened ellipsoid with planar opposing faces (~250 Å × ~240 Å × ~166 Å). The shape is oval when viewed from the top. The planar top and bottom surfaces are parallel and ~166 Å apart. Lateral views of LDL reveal an approximately rounded trapezoid with grooves on the front and back surfaces (Fig. 1C bottom).

Notably, the other ~40,000 particle images are qualitatively similar to the selected subpopulation, but tend to be somewhat larger or rounder in appearance. We also observed larger particles with additional striations. While the protein content of LDL is fixed, the amount of triglyceride, lipids, and cholesterol varies. Nothing in the images of these other particles invalidates our analysis of the subpopulation, but due to the large variability within the set, and the use of one orientation in the imaging, we cannot produce other reliable three-dimensional structures from these data. Proper characterization of the heterogeneity within this population may use a technique like electron cryomicroscopy followed by posttomographic classification, alignment, and averaging (10).

The features observed in raw micrographs of the LDL · LDLr complex are essentially identical to those of LDL, except for the appearance of an additional high-density protrusion on the surface (Fig. 1D, E). The reconstructed map of the LDL · LDLr complex is also similar to that of LDL, containing the same shape and dimensions as LDL with an additional ~35–45 Å protrusion on the surface (Fig. 1F). The size and location of this protrusion corresponds well to the additional high-density region observed in raw micrographs of the LDL · LDLr complex (Fig. 1E left). The dimensions of the protrusion determined from a difference map (Fig. 2, gold mesh) is most compatible with that of the LDL receptor β-propeller with a part of EGF-like repeat C domain, a ~40 kDa domain with a diameter of ~40–50 Å as seen in its crystal structure (11) (Fig. 2, red). Though the LDLr location is clear, its orientation is ambiguous due to the roundness of the density. In contrast, the cysteine-rich repeats and EGF-like repeats are much smaller (11 to 15 Å each based on α-helix distances). Moreover, the cysteine-rich repeats or EGF-like repeats are unlikely to interact with each other to form a larger module because at neutral pH the LDLr extracellular domain is elongated (11–13) and at acidic pH these modules interact with each other only via their linker regions (11). While LDL binding to LDLr certainly in...
volves cysteine-rich repeats of the ligand binding domain because their deletion reduces binding to <20% of the wild-type LDLr, two studies support the hypothesis that at neutral pH the β-propeller also interacts with lipoproteins, because receptors with amino acid substitutions in the β-propeller bind less ligand than wild-type LDLr (13, 14).

The internal density distribution of LDL shows an organized core. To examine the internal core, density maps of LDL and LDL · LDLr were cut into halves along two orthogonal directions perpendicular to the plane of the top surface (Fig. 3, A and C). By contouring the cross sections from low-density (3.04σ, in blue) to high-density (6.84σ, in red) using a color ramp scheme, the cut-away surfaces display three flat, internal higher-density regions (green striations), which are parallel to each other and to the highest density planar top surfaces (yellow, orange, and red). Much lower density and chambers (blue and gray holes) appear on both sides of these high-density internal striations. The three internal high densities form what are roughly three-layers of “isthmi” with similar dimensions. The density near the center of each “isthmus” is stronger than that near the edges. The isthmi are parallel to each other and separated by ∼35 Å. The two outer isthmi are also parallel to the outer shell of LDL, at a distance of ∼31 Å.

We propose a liquid crystalline core model of LDL in which the CE molecules are arranged in stacks with their sterol moieties side-by-side in the higher-density regions while the fatty acyl chains extend from either side, and are observed as parallel lower-density compartments (Fig. 3B, D). Indeed our model builds on the early report of LDL as a discoidal, i.e., non spherical particle with apoB in two ring-shaped structures around the perimeter (15). Later studies using image reconstruction showed LDL as more spheroidal than discoidal and the presence of striations (16) but with less detail than in the present study where we observe a non uniform distribution of high and low density that we assign to protein- and lipid-rich surface regions respectively. Unlike the earlier structure (16), in the present study the
The density of LDL, 1.03–1.06 g/ml is derived from the weighted average of its high-density component, protein, and the low-density component, lipids (17). Phospholipids and protein each contribute about half of the surface components of LDL. Thus, the LDL density map is expected to contain regions of different densities that correspond to the apo B-100 and the lipids respectively. To reveal the high density regions in the three-dimensional map of LDL, we display the map at two contour levels, 5.67σ (yellow solid) and 4.76σ (green mesh; Fig. 4A). At a higher contour level of 5.67σ, we assign the corresponding region to be that of apo B-100, e.g. ~18% of entire LDL volume. At a slightly lower contour level of 4.76σ, the map volume is about twice that of the apo B-100 volume and is a reasonable estimation of apo B-100 with bound phospholipids. These two contour levels together reveal an apo B-100 configuration that appears as a pair of paddles (Fig. 4A, top and bottom faces), connected at one end of the particle by a linker region (Fig. 4A, right and left faces), with three separate semicircular and long “fingers” extending from each edge of the linker region to wrap around the particle (Fig. 4A, front and back faces). These fingers have similar lengths (~230 Å) and are separated by at least ~35 Å. Notably, the finger-shaped densities on the “front” are weaker than on the back suggesting that the front region is lipid-rich and distinct from the rest of the particle surface. Thus, the overall shape of the high density surface features suggests that the apo B-100 molecule comprises top and bottom paddles that are attached by a linking region, with a railing of semicircular fingers on each side.

According to sequence-based structural predictions, apo B-100 can be divided into five domains, including two amphipathic β-sheet rich (β₁, β₂), two amphipathic α-helix rich (α₂, α₃), and a mixed α-β domain (βα) (Fig. 4D). Each α-helix-rich domain comprises more than 500 amino acid residues that are predicted to form 4–5 unusually long single α-helices (80–240 Å each, Fig. S3). The β-sheet rich domains are predicted to have more than ~80% of their residues forming β-strands or loops. Similar analysis led previously to the so-called pentapartite model of apo B-100 (20), but the spatial distributions of these five domains on the LDL particle have remained unassigned. In general, a low resolution map cannot resolve the structural motifs of a protein. However in our case, we can assign the spatially discrete high density regions of the map to domains in the pentapartite model, i.e. the α-helix rich and β-sheet rich domains by considering: (i) the unusual structural motifs of apo B-100 with several unusually long α-helices in the two α-helix rich domains, (ii) the consensus that the apo B-100 is distributed throughout the surface of the particle (5, 7, 9), and finally (iii) the scattering density difference between protein and lipids, (Fig. 4B, C). In our model, the paddles (top and bottom faces) and linker (right face) comprise the amphipathic β-sheet domains, and the railings of semicircular fingers (front and back faces) contain the amphipathic α-helices. The continuous amphipathic β-sheets domains, β₁ and β₂, are likely anchored to the core lipids via their hydrophobic faces, with few if any interspersed phospholipids. The more flexible α₂ and α₃ domains containing multiple stretches of α-helices likely orient their amphipathic helical axes coplanar to the surface monolayer of phospholipids as seen in the front and back of the particle. The stacks of acyl chains in the CE core are directed outward towards the amphipathic β-sheets domains on the top and bottom faces of the particle (Fig. 4B and D) and are semicircular surrounded by the flexible amphipathic α-helix rich domains (Fig. 4E).

According to this model and the reconstructed maps (Figs. 3 and 4), the LDL β-propeller domain binds to the linker between the amphipathic β-sheets domains that comprise the top and bottom paddles (Fig. 4A, right, asterisk). This assignment is consistent with the biochemical data of others. The apoB antibodies 4G3, 278, 3G9, and 5E11 which block LDL binding to its receptor (21, 22) are known to bind apoB residues in the region containing amino acids 3000–3500, a region that the pentapartite model of LDL
apoB predicts to be a β-sheet domain. The cryoEM reconstruction of the LDL · LDLr complex localizes the β-propeller to a β-sheet domain (β₁ or β₂) of the pentapartite model of LDL as well (Fig 4, asterisk). Our finding is thus consistent with the biochemical studies using antibodies. However, since the cryoEM map is low-resolution, we cannot with confidence discriminate between β₁ and β₂ in assigning their locations to specific densities.

Our model reveals not only protein-rich regions but also protein-poor, lipid-rich regions (Fig. 4 left face) that are potential sites where enzymes and lipid transfer proteins could interact with the phospholipid surface via their putative interfacial recognition sites. The most notable of these is cholesteryl ester transfer protein, which transports CE from HDL to LDL and very low-density lipoprotein (VLDL), and is known to bind to phospholipid surfaces that contain no protein (23).

Our structure was obtained after quenching from below the LDL thermal transition temperature so that most particles were captured in a liquid crystalline phase. A similar study with LDL quenched from 37 °C would likely produce fewer liquid crystalline LDL particles. Nevertheless, our structure is likely relevant to LDL physiology. The major determinant of the transition temperature is the triglyceride content, which is lowest (4.5 ± 0.4%) for the middle density fraction (2), which was used in our study. Because of the breadth of the transition there is still a fraction of LDL that is liquid crystalline even at 37 °C. In addition, in the liver the triglyceride content of LDL could be further lowered by hepatic lipase (24) thereby increasing the number of liquid

Fig. 4. High density outer features of LDL reveal the structure of apoB-100. (A) Six orthogonal views are shown. The high density region of the LDL three-dimensional map is displayed at the contour level of 5.67σ (yellow) and 4.76σ (green mesh). The lowest density features of the outer shell map to the ‘front’ and ‘left’ views. The asterisk in the ‘right’ view indicates the location of the putative β-propeller of the LDLr. (B) Six orthogonal projection views of the pentapartite model of apo B-100 (two amphipathic β-sheet rich domains interrupted by two amphipathic α-helix-rich domains, shown in yellow cable) onto the three-dimensional reconstruction of LDL. (C) Superposition of the high outer shell density (beige translucent solid), i.e., apo B-100 from (A) and the domain structure of the pentapartite model (yellow cable) from (B). (D) Probability (P = 0, low; P = 1, high) of helical (red, Pα) and β-strand (blue, Pβ) regions in apo B-100 based on sequence prediction suggests a five domain model designated as α₂, α₃, β₁, β₂ and βα₁. (E) A cut-away view of the LDL model shows the surface structure of apo B-100 and the internal organization of the CE sterol moieties. In (B), (C), and (E), the phospholipid head groups, CE, and TG are displayed as cyan, magenta, and blue balls respectively.
crystalline particles in vivo. Finally, the surface densities defined by the distribution of protein and lipid in LDL particles containing an isotropic liquid core are likely to share some similarities with our structure, although this will have to be experimentally verified.

Our model of apo B-100 and CE packing is consistent with many biochemical and biophysical measurements (5, 6, 16–19, 21–23, 25, 26), and divides apo B-100 into multiple domains that surround a core with a distinctive liquid crystalline structure. The structural organization between CE and apo B-100 may bear relevance to human pathobiology. While normal individuals have a predominance of large LDL particles, some patients have an atherogenic lipoprotein phenotype that is characterized by small, dense LDL particles; this phenotype is associated with low plasma HDL-cholesterol, insulin resistance, and increased risk for coronary heart disease (27). Unlike the core of normal LDL which is liquid crystalline (Fig. 3), the small, dense LDL is triglyceride-enriched and has an isotropic liquid core with no apparent striations according to cryoEM (18). Furthermore, the apo B-100 of triglyceride-rich LDL also shows impaired binding to fibroblast LDL receptors, suggesting an altered conformation (28, 29).

Coronary heart disease (27) is associated with impaired cellular LDL uptake.

CryoEM Data Collection. Micrographs of LDL and LDL · LDLr particles were acquired at liquid helium temperature (4−25 ° Kelvin) using a JEM30000SF electron cryomicroscope (JOEL) with a field emission gun operated at 300 kV high tensions. The microscope was equipped with a Gatan US4000 4k × 4k CCD camera (Gatan) with a condenser aperture of 100 μm, a spot size of three and an objective lens aperture of 70 μm. Images of LDL were recorded with a microscope magnification of ∼40,000x, a dose of ∼20 e−/Å², and a defocus of 1.5 to 3.5 μm; those of LDL · LDLr complex were recorded similarly but with a Gatan 2K × 2K CCD camera at a microscope magnification of 50,000x.

Image Processing. Individual particle images were selected using the program boxer in the EMAN package (30). A total of 47,593 particles were picked automatically and manually checked to remove overlapped particles or damaged particles from 654 LDL CFD frames with a 112 × 112 pixel box (−476 Å × 476 Å at the specimen); 12,547 particles were boxed from 275 LDL · LDLr complex CFD frames, using a 192 × 192 pixel box (−518.4 Å × 518.4 Å at the specimen). CTF and envelope-function parameters were fitted using the automated fitting program fitctf,py and subsequently manually adjusted by ctffit (EMAN suite; 30). Applyctf (EMAN suite) was used to apply phase corrections to the particles and store CTF parameters for later amplitude correction. For LDL and the LDL · LDLr complex, 1,969 and 2,168 class averages were used respectively.

Three-Dimensional Reconstruction. To prevent bias from a starting model, a smooth 297 Å × 270 Å × 243 Å Gaussian blob was generated by makeini- tialmodel.py and used for refinement during reconstruction of both LDL and the LDL · LDLr complex models. For the beginning eight refinement rounds, only very low-resolution information for the particles was used (below the first CTF zero in reciprocal space). Iterative refinement proceeded until convergence. For higher resolution refinement, CTF amplitude correction was performed, finer angular sampling was used, and solvent flattening via molecular mas was imposed. This process was iterated until convergence as well. According to the 0.5 Fourier shell correlation criterion (31), the final resolutions of the asymmetric reconstructions for LDL and the LDL · LDLr complex were −28 Å.

Structural Analysis. Visualization and segmentation of the three-dimensional reconstructions were performed using UCSF Chimera (32). The LDL maps were analyzed at different contour levels based on the estimated volumes of the total LDL particle (3.87 × 10⁵ Å³) and the volume of apo B-100 (6.46 × 10⁵ Å³). To analyze the LDL · LDLr complex, the final LDL reconstruction was scaled and aligned to the final LDL · LDLr complex reconstruction and a difference map calculated using proc3d (30). To locate the LDLr on the LDL surface, the β-propeller domain with a part of EGF-like repeat C domain from the crystal structure of the extracellular domain of the LDLr (PDB: 4KX2, residues 601–881) was used. A total of 325 residues or 2,450 atoms (11) was automatically docked to the difference map between the LDL · LDLr complex and LDL, and contoured at 1.3σ. The protein secondary structure predictive algorithm, Predictprotein (33), was used to determine the probability of apo B-100 regions forming an α-helix (Pₐ) or β-strand (Pₜ). The distribution of these regions of secondary structure has been used to generate a pentapartite model of apo B-100.

ACKNOWLEDGMENTS. Ren thanks D.A. Agard, K.H. Weisgraber at UCSF and M.F. Schmid at Baylor College of Medicine for discussions. This project has been supported by the Biomedical Technology Research Center for Structural Biology (P41RR02250) and P01GM099116; Rüdenko was supported by an American Heart Scientist Development Award. Ren was partially supported by W.M. Keck Foundation. Pownall was supported by HL-30914 and –56865.

Fig. S1. Single-particle image reconstruction of LDL and LDL·LDLr complexes. Selected reference projections of the final three-dimensional reconstruction (A, left, projection) versus corresponding class averages (B, right, average) of LDL (left) and LDL·LDLr (right).
Fig. S2. Single-particle image reconstruction of LDL and LDL · LDLr complexes. (A) Histogram of the raw images in each view of the final model showing that all views of the LDL (left) and LDL · LDLr (right) complex are well represented and that there are no obvious preferential orientations. Images were aligned translationally (x and y) and rotationally in-plane (ω), and each particle was assigned to the group to which it correlated best. Each circle represents a projection of the final model for a given Θ and Φ value. (B) Fourier shell correlation to determine map resolution. The total dataset of images was divided into two equal groups, the particles in each dataset were aligned independently with projections of the final model, and two independent maps were correlated in Fourier space. The curves cross the 50% threshold at 28 Å resolution for LDL and LDL · LDLr reconstructions.
Fig. S3. Sequence-based structure prediction of two α-helix rich domains, \( \alpha_2 \) and \( \alpha_3 \). (A) The analysis suggests the \( \alpha_2 \) domain contains six unusually long helices (highlighted in red), which are \( \approx 90 \text{ Å} \) long (60 a.a.), \( \approx 78 \text{ Å} \) (52 a.a.), \( \approx 75 \text{ Å} \) (50 a.a.), \( \approx 237 \text{ Å} \) (158 a.a.), \( \approx 35 \text{ Å} \) (23 a.a.), and \( \approx 63 \text{ Å} \) (95 a.a.), respectively. In total, 438 out of 520 a.a. (\( \approx 90\% \)) are involved in alpha helical structure. (B) Similarly, the \( \alpha_3 \) domain contains 8 α-helices. The length of the helices are about \( \approx 65\% \) of the 500 a.a. residues comprising the \( \alpha_3 \) domain can be assigned to α-helices. The short helices are mostly near the beginning or ending portion of both α-domains which are highlighted in gray and green. The small loops between long helices are highlighted in purple.