Folding, Calcium Binding, and Structural Characterization of a Concatemer of the First and Second Ligand-Binding Modules of the Low-Density Lipoprotein Receptor†

Stephan Bieri, Annette R. Atkins, Huang T. Lee, Donald J. Winzor, Ross Smith, and Paulus A. Kroon*

Centre for Protein Structure, Function, and Engineering, Department of Biochemistry, University of Queensland, Brisbane, Queensland 4072, Australia

Received February 25, 1998; Revised Manuscript Received June 8, 1998

ABSTRACT: The ligand-binding domain of the low-density lipoprotein (LDL) receptor is comprised of seven tandemly repeated ligand-binding modules, each being approximately 40 amino acids long and containing six conserved cysteine residues. We have expressed and characterized a concatemer of the first two modules (LB_1 and LB_2) of the human LDL receptor. Oxidative folding of the recombinant concatemer (rLB_{1-2}), in the presence of calcium ions, gave a single dominant isomer with six disulfide bonds. Peptic cleavage of the short linker region that connects the last cysteine residue of LB_1 and the first cysteine residue of LB_2 yielded two discrete fragments, thus excluding the presence of intermodule disulfide bonds. The N-terminal module, LB_1, reacted with a conformation-specific monoclonal antibody (IgG-C7) made to LB_1 in the native LDL receptor. From this, we concluded that the first module was correctly folded, with the same set of disulfide bonds as LB_1 of the LDL receptor. The disulfide bond connections of LB_2 were identified from mass spectral analysis of fragments formed by digestion of the C-terminal peptic fragment with elastase. These data showed that the disulfide bonds of LB_2 connected Cys(I) and Cys(III), Cys(II) and Cys(V), and Cys(IV) and Cys(VI). This pattern is identical to that found for recombinant LB_1 and LB_2. The concatemer has two high-affinity calcium-binding sites, one per module. An analysis of the secondary chemical shifts of Cα protons shows that the conformations of LB_1 and LB_2 in the concatemer are very similar to those of the individual modules, with no evidence for strong interactions between the two modules.

The low-density lipoprotein (LDL)[1] receptor (LDLR) is a transmembrane glycoprotein that removes cholesterol-rich lipoproteins from circulation by receptor-mediated endocytosis (*1*, *2*). Mutations in the LDLR that interfere with receptor-mediated endocytosis of lipoproteins lead to the genetic disease familial hypercholesterolemia that is characterized by elevated plasma cholesterol levels, accelerated atherosclerosis, and premature coronary artery disease (*3*).

The LDLR is the prototype of the multidomain LDLR superfamily. All members of this gene family have an extracellular ligand-binding domain that consists of one or more strings of tandemly repeated ligand-binding modules (*4*). These modules, which are also referred to as type A modules, are approximately 40 amino acids long and contain six highly conserved cysteine residues that form three intramodule disulfide bonds (*5*, *6*). The ligand-binding domain of the LDLR consists of a single string of seven ligand-binding modules. This combination of modules allows the LDLR to bind plasma lipoproteins that contain apolipoprotein (apo) B-100, a 550 kDa glycoprotein which is present as a single copy on lipoprotein particles (*2*), and apoE, a 34 kDa glycoprotein which is present as multiple copies on lipoproteins (*7*).

We have previously characterized the first two ligand-binding modules (LB_1 and LB_2) of the human LDLR as autonomously folding domains that contain three disulfide bonds with a Cys(I)−Cys(III), Cys(II)−Cys(V), and Cys(IV)−Cys(VI) connectivity (*5*, *6*). Both modules were expressed as recombinant peptides (rLB_1 and rLB_2) and in the presence of calcium had a well-defined β-hairpin structure followed by several turns (*8*, *9*). Recombinant ligand-binding modules have provided a powerful tool for the elucidation of the molecular basis for defects caused by mutations in ligand-binding modules (*10*, *11*).

The LDLR uses different combinations of its string of seven ligand-binding modules to bind apoB-100 and apoE (*12*, *13*). As a first step toward understanding the structure of a string of ligand-binding modules, we have expressed and characterized a concatemer of LB_1 and LB_2 (rLB_{1-2}). This recombinant peptide contains 12 cysteine residues

† This work was supported by a research grant from the National Health and Medical Research Council of Australia.

* Address correspondence to P. A. Kroon, Department of Biochemistry, University of Queensland, Brisbane, Queensland 4072, Australia. Telephone: +61-7-3365-4876. Fax: +61-7-3365-4699. E-mail: pkroon@biosci.uq.edu.au.

[1] Abbreviations: LDL, low-density lipoprotein; LDLR, LDL receptor; LB_1 and LB_2, first and second ligand binding modules, respectively, of the human LDL receptor; rLB_1 and rLB_2, recombinant LB_1 and LB_2, respectively; LB_{1-2}, concatemer of the two amino-terminal ligand-binding modules; rLB_{1-2}, recombinant LB_{1-2}; $LB_1(45)$, N-terminal peptic fragment of rLB_{1-2}; $LB_2(40)$, C-terminal peptic fragment of rLB_{1-2}; apo, apolipoprotein; RP, reverse-phase; GdnHCl, guanidine hydrochloride; CD, circular dichroism; ER, endoplasmic reticulum.

S0006-2960(98)00452-8 CCC: $15.00
Published on Web 07/15/1998

leading to a possible 10 395 disulfide bond pairings. We show that rLB_{1-2} folds autonomously in the presence of calcium ions into a single isomer, comprised of two discrete ligand-binding modules with a characteristic 1−3, 2−5, and 4−6 disulfide-bonding pattern. The concatemer requires calcium for folding and in the folded state has two high-affinity calcium-binding sites. An analysis of the secondary chemical shifts of Cα protons shows that the conformations of LB_1 and LB_2 in the concatemer are very similar to those of the individual modules, with no evidence for strong interactions between the two modules.

MATERIALS AND METHODS

Expression and Folding. The cDNA encoding the first and second cysteine-rich modules of the ligand-binding domain of the human LDLR (residues 1−83) was cloned into the expression vector pGEX-2T (Pharmacia), to give pGST-LB_{1-2}. This vector encodes a thrombin-cleavable fusion protein of GST and rLB_{1-2}. The fusion protein was expressed in *Escherichia coli* DH5α cells, purified by affinity chromatography on glutathione−agarose, and cleaved with thrombin as decribed previously (*5*, *6*). The purified peptide was folded overnight, in the presence of 3 mM GSH/0.3 mM GSSG in 50 mM Tris-HCl (pH 8.5), 150 mM NaCl, and 2.5 mM $CaCl_2$ at 4 °C. rLB_1 and rLB_2 were expressed, purified, and folded as described previously (*5*, *6*).

To assess the requirement for calcium in the folding reaction, rLB_{1-2}, in 20 mM NH_4HCO_3 (pH 8.5), was reduced with 45 mM DTT for 15 min at 50 °C. The reduced peptide (2 μg) was purified by RP-HPLC, dried in a Speed Vac Concentrator, and resuspended in 50 mM Tris-HCl (pH 7.5) and 150 mM NaCl that contained 2.5 mM $CaCl_2$, 2.5 mM Na_2EDTA, or 2.5 mM $MgCl_2$. Reduced and oxidized glutathione were added to final concentrations of 3 and 0.3 mM, respectively, and the pH was adjusted to 8.5. The samples were incubated for various times at 4 °C and analyzed by RP-FPLC.

HPLC and FPLC Analyses. Refolded rLB_{1-2} was purified using an Econosil C18 reverse-phase column (10 μm particle size, 60 μm pore, 10 mm × 250 mm) obtained from Alltech (Deerfield, IL). A trifluoroacetic acid/acetonitrile solvent system was used (solvent A, 0.1% trifluoroacetic acid in water; solvent B, 0.1% trifluoroacetic acid in acetonitrile) for the purification, with an 18 to 40% trifluoroacetic acid/acetonitrile gradient, over the course of 44 min, with a flow rate of 2.5 mL/min.

FPLC analyses were performed using a Smart System (Pharmacia). A μRPC C2/C18 reverse-phase column (3 μm particle size, 120 Å pore, 2.1 mm × 100 mm) and a trifluoroacetic acid/acetonitrile buffer system (see above) were used for all FPLC analyses.

Mass Spectrometry. Peptides were analyzed using a triple-quadrupole mass spectrometer (SCIEX API III, Perkin-Elmer, Thornhill, Canada), equipped with an ion-spray interface and operated in the positive detection mode. Mass spectra were analyzed using the MacSpec version 3.2 software (Perkin-Elmer). LC−MS analyses were performed using Applied Biosystems 140B dual-syringe pumps and a C18 RP column (10 μm particle size, 300 Å pore, 4.6 mm × 250 mm). Samples were eluted using a 1 to 40% trifluoroacetic acid/acetonitrile gradient over the course of 40 min.

Reduction and Alkylation. To alkylate free cysteine thiol groups on rLB_{1-2}, the peptides were denatured in the presence of 6 M guanidine hydrochloride (GdnHCl) for 2 h at 50 °C and immediately alkylated with supersaturated iodoacetamide (approximately 2.2 M) as previously described (*5*). The samples were immediately injected onto a C2/C18 RP-FPLC column on a Smart System. The column was washed with 0.1% trifluoroacetic acid/10% acetonitrile until the effluent absorbance at 214 nm dropped below 0.05, and the peptide was eluted as described above.

Molar Extinction Coefficient. The molar extinction coefficient for rLB_{1-2} ($\epsilon_{280} = 14\,427\ M^{-1}\ cm^{-1}$) was measured experimentally, by subjecting a sample with a known absorbance at 280 nm to quantitative amino acid analysis using PicoTag reagents (Waters).

Proteolytic Digestions. RP-HPLC-purified peptide (2−5 μg) was dried in a Speed Vac Concentrator and resuspended in 10 mM HCl for digestion with pepsin or in 20 mM NH_4HCO_3 (pH 8.5) for digestion with elastase. Digestions with pepsin and elastase were carried out with a protein:enzyme ratio of 10:1, for 16 h at 37 °C.

Immunoblot Analysis. Production of the structure-specific monoclonal antibody, IgG-C7 (American Type Culture Collection, Rockville, MD), and the immunoblot analysis were carried out as previously described (*5*).

Zonal Chromatography. The binding affinities of calcium for rLB_{1-2} were determined by an FPLC adaptation of the procedure of Hummel and Dreyer (*14*) on a Smart System. Buffers were treated with Chelex-100 (Bio-Rad) to remove traces of heavy metal ions. A PC3.2/10 fast desalting (Sephadex G-25 resin) column (Pharmacia) was equilibrated with 50 mM Tris-HCl (pH 7.5) and 0.15 M NaCl that contained various concentrations of $CaCl_2$ (5−364 μM, 0.5 nCi/μL of $^{45}CaCl_2$). A 100 μL volume of rLB_{1-2} (∼90 μM) in equilibration buffer was applied to the column and eluted at 20 °C with a constant flow rate of 200 μL/min. The elution profile was monitored spectrophotometrically at 280 nm, and individual fractions of 100 μL were collected from 1 min prior to injection up to 8 min after injection. The amount of peptide eluted was calculated from the area under the A_{280} curve using the experimentally determined ϵ_{280} of $14\,427\ M^{-1}\ cm^{-1}$. An aliquot (50 μL) from each fraction was counted in a Beckman LS3801 scintillation counter. The concentration of calcium in individual fractions was calculated from the measured counts.

Sedimentation Analysis. To meet the requirements for dialysis equilibrium in samples subjected to sedimentation equilibrium (*15*), samples (∼250 μL) were loaded onto a 0.9 mL Sephadex G-15 column, equilibrated with 50 mM Tris-HCl (pH 7.5), 0.15 M NaCl, and either 0.45 mM $CaCl_2$, 10 mM $CaCl_2$, or 30 mM Na_2EDTA (equilibration buffer), and were eluted with 350 μL of the same buffer. The final concentrations of the peptide were determined spectrophotometrically at 280 nm.

Peptides were loaded into the sample sector of a double-sector cell, with equilibration buffer in the reference sector. Cells (path lengths of 1.5, 2.5, or 11 mm) were chosen to keep the A_{280} at 1.0 or lower. Samples were centrifuged at 35 000 rpm and 20 °C in a Beckman Optima XL-A analytical ultracentrifuge using a Beckman An Ti 60 rotor. After 9 h, the sedimentation distributions were scanned three times at

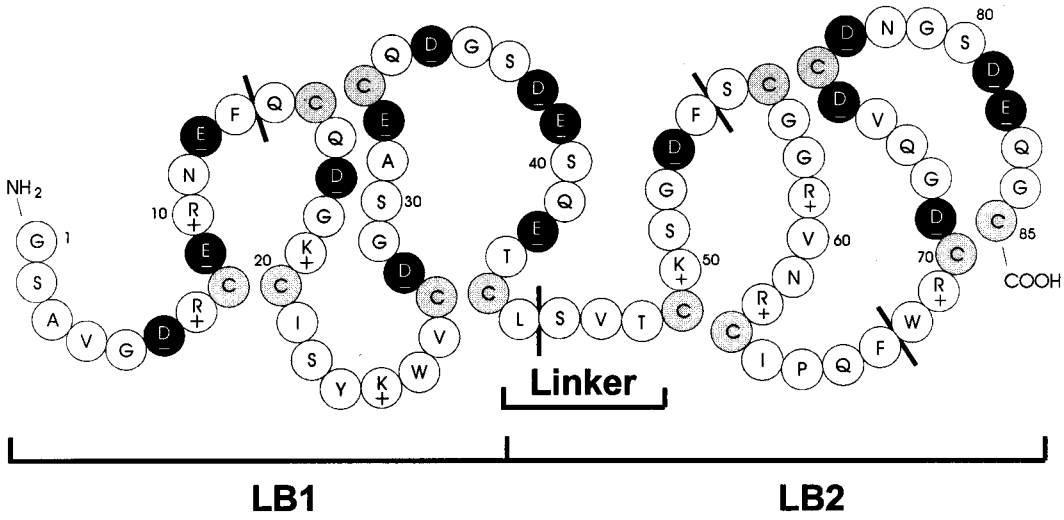


FIGURE 1: Amino acid sequence of rLB_{1-2}. The sequence contains the first 83 amino acids of the mature LDLR with an N-terminal Gly-Ser extension. Charged residues are labeled with a + or −. Acidic residues are shown in black and cysteine residues in gray. Predicted pepsin cleavage sites are indicated by a line. Cysteine residues that are disulfide-bonded in rLB_1 and rLB_2 are located next to each other. The four-amino acid linker that joins the last cysteine residue of LB_1 with the first cysteine residue of LB_2 is labeled.

280 nm at 3 h intervals until no time dependence between successive scans could be detected.

Partial specific volumes (ν) were estimated from the amino acid composition as described (*16*, *17*). Solvent densities (ρ) were calculated as described (*16*). Data were analyzed using software provided by G. Ralston (Department of Biochemistry, University of Sydney, Sydney, Australia).

NMR Analyses. ^1H NMR spectra were recorded on a 1.9 mM solution of rLB_{1-2} in 10% ^2H_2O on a Bruker DRX 750 MHz spectrometer. $CaCl_2$ (15 mM) was added to provide an approximate 4-fold molar excess of calcium, assuming each module binds a single calcium ion. TOCSY (*18*) and NOESY (*19*) spectra were recorded at two temperatures (300 and 310 K) and two pH values (pH 5.1 and 6.1) to facilitate sequence-specific assignments. An MLEV-17 spin lock sequence of either 63 or 82 ms was employed in the TOCSY experiments, and a 200 ms mixing time was used when collecting NOESY spectra. Time-proportional phase incrementation was utilized for F_1 quadrature detection, and water suppression was accomplished with the WATERGATE sequence (*20*). Typically, spectra were recorded as 470−512 t_1 increments, collected into 4096 data points, and apodized using a 90°-shifted sine-squared function in both dimensions within the Bruker software XWINNMR. Chemical shifts are referenced to the β-methyl group of Ala3 as measured in rLB_1 (1.41 ppm), assuming that the conformation of the N-terminal residues is unaffected by the presence of the second module. The CαH secondary chemical shifts, defined as the difference between the observed and random coil values (*21*), were determined using the random coil shifts of Merutka (*22*).

Circular Dichroism Spectroscopy. Circular dichroism (CD) spectra were recorded on a Jasco J-710 spectropolarimeter (Tokyo, Japan), over the wavelength range of 260−188 nm, using a 0.1 mm path length cell, a bandwidth of 1.0 nm, a response time of 2 s, and a scan rate of 20 nm min^{-1}. Several scans were averaged. For these analyses, rLB_{1-2} was dissolved in 50 mM Tris-HCl (pH 7.5) and 0.15 M NaCl at a final concentration of 45 μM, and with variable concentrations of $CaCl_2$.

RESULTS

Expression and Calcium-Dependent Folding. The concatemer of LB_1 and LB_2 was expressed in *E. coli* as a thrombin-cleavable GST fusion protein, using the expression vector pGST-LB_{1-2}. This vector encodes GST, followed by the first 83 amino acids of the mature human LDLR. Cleavage with thrombin releases an 85-amino acid peptide comprised of LB_1 (residues 3−44) and LB_2 (residues 45−85) (*4*), with an N-terminal Gly-Ser extension derived from the thrombin cleavage site. This peptide is illustrated in Figure 1.

Theoretically, a peptide with 12 cysteine residues can form 10 395 distinct isomers, each with six disulfide bonds. When fully reduced rLB_{1-2} was incubated in the presence of reduced and oxidized glutathione and 2.5 mM calcium ions for 6 h, most of the peptide folded into a single form after 6 h, and folding was complete after 48 h. In contrast, when calcium was not included, reduced rLB_{1-2} folded into a complex mixture of isomers, without any preference for the correctly folded form. Attempts to fold rLB_{1-2} in the presence of Mg^{2+} led to the formation of multiple oxidized species, again with no evidence for the correctly folded form. These data show that calcium is required for folding, and that the requirement for calcium is specific.

The refolded isomer of rLB_{1-2} was purified by RP-HPLC, and used for subsequent analyses.

The Folded Isomer of rLB_{1-2} Is Fully Oxidized. Mass spectral analysis of RP-HPLC-purified rLB_{1-2} gave a molecular weight of 9238 ± 1. A comparison with calculated molecular weights of fully reduced (9250) and oxidized (9238) forms suggests that the refolded isomer is fully oxidized. To confirm the complete absence of free cysteine thiol groups, rLB_{1-2} was incubated in the presence of 6 M GdnHCl for 2 h at 50 °C to expose thiol groups which might be buried. Supersaturated iodoacetamide was subsequently added to alkylate these groups. The data in Figure 2A show that the molecular weight of rLB_{1-2} was unchanged (9238), with no evidence for species with alkylated cysteine residues. In contrast, when rLB_{1-2} was first incubated with DTT, the molecular weight was increased to 9935 (Figure 2B),

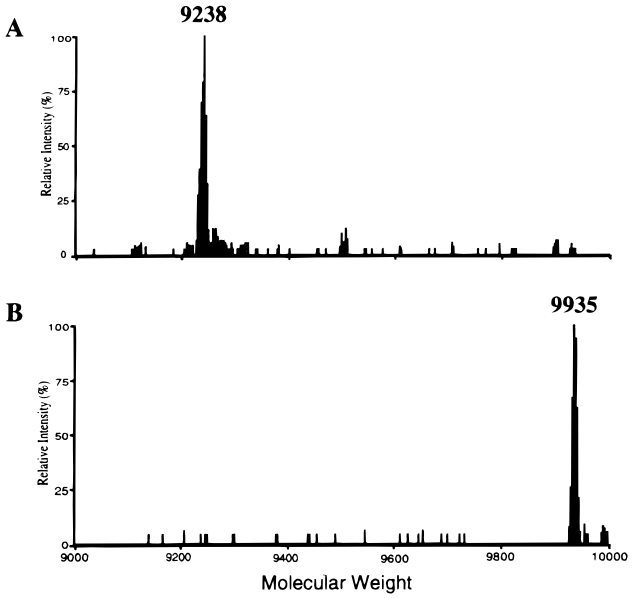


FIGURE 2: Mass spectral analysis of rLB_{1-2} for the presence of alkylated thiol groups. rLB_{1-2} was incubated in 6 M GdnHCl for 2 h at 50 °C in the absence (A) and presence (B) of 45 mM DTT and injected into supersaturated iodoacetamide as described in Materials and Methods. The treated peptides were immediately purified by RP-FPLC and analyzed by mass spectrometry.

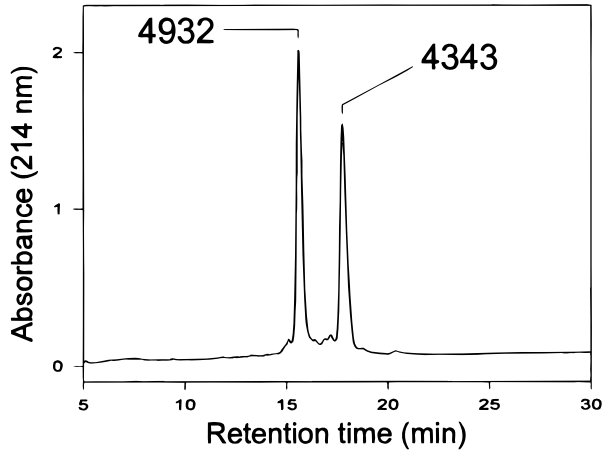


FIGURE 3: RP-FPLC analysis of rLB_{1-2} cleaved with pepsin. Peptic cleavage fragments were separated by RP-FPLC and analyzed by electrospray mass spectrometry.

consistent with the presence of 12 alkylated cysteine residues. These data show that the refolded isomer of rLB_{1-2} was completely oxidized and contained no free cysteine thiol groups.

The Folded Isomer of rLB_{1-2} Contains No Intermodule Disulfide Bonds. To determine whether intermodule disulfide bonds existed, we cleaved the peptide within the short linker region (LSVT) that joins the last cysteine residue of LB_1 and the first cysteine residue of LB_2. For these experiments, we used pepsin. This proteolytic enzyme cleaves Leu−X and Phe−X bonds and is therefore predicted to cleave the linker between the leucine and serine residues. This and all other potential pepsin cleavage sites in rLB_{1-2} are labeled with a line in Figure 1.

Pepsin cleaved rLB_{1-2} quantitatively into two fragments (molecular weights of 4932 and 4343) that were clearly resolved by RP-FPLC (Figure 3). The molecular weight of the first fragment corresponded to that calculated for the amino-terminal cleavage fragment of rLB_{1-2} (residues 1−45), with three disulfide bonds. We refer to this peptide as LB_1-(45). The second fragment had a molecular weight that was 18 mass units higher than that calculated (4325) for the C-terminal fragment of rLB_{1-2} (residues 46−85), again with three disulfide bonds. We refer to the C-terminal peptide as LB_2(40).

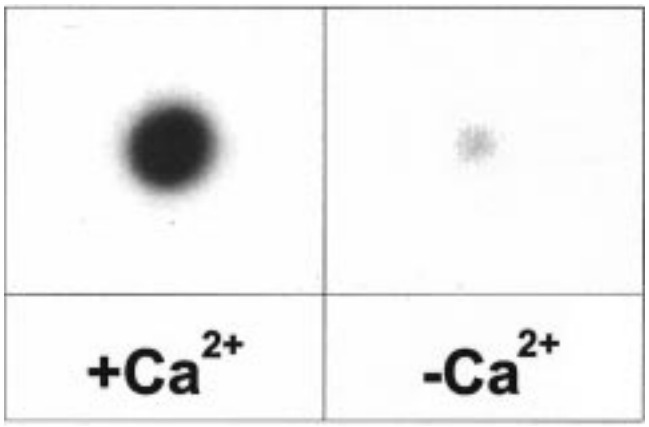
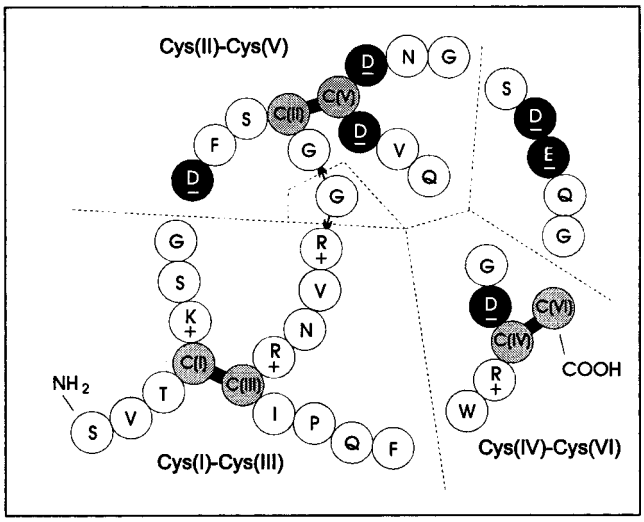


FIGURE 4: Dot blot analysis of rLB_{1-2}. HPLC-purified rLB_{1-2} (4 μg/dot) was dotted onto a 0.2 μm nitrocellulose membrane. The membrane was baked overnight in a vacuum oven, at 80 °C, and blocked for 1 h in the presence of 5%, nonfat skim milk. Membranes were incubated with IgG-C7 in the presence (+) or in the absence (−) of 2 mM Ca^{2+}. Incubation mixtures that lacked calcium contained 30 mM EDTA. Signals were detected by enhanced chemiluminescence followed by exposure to X-ray film.

The increase of 18 mass units suggested that LB_2(40) had been cleaved internally at either Phe54 or Phe67 (Figure 1). To confirm this, LB_2(40) was reduced with DTT and free cysteine residues carboxymethylated with iodoacetamide. RP-FPLC analysis of the reaction mixture yielded two fragments with molecular weights of 2533 and 2158 (data not shown). These molecular weights are predicted for carboxymethylated fragments Ser46−Phe67 and Trp68−Cys85, consistent with cleavage between Phe67 and Trp68.

Pepsin therefore cleaves rLB_{1-2} at two (Leu45 and Phe67) of the four possible pepsin sites (Figure 1). The quantitative conversion of rLB_{1-2} into N-terminal [LB_1(45)] and C-terminal [LB_2(40)] fragments is only possible if both modules contain three intramodule disulfide bonds, and rules out the possible existence of intermodule disulfide bonds.

Disulfide Bonds in LB_1. We have previously used a conformation-specific monoclonal antibody, IgG-C7, to show that rLB_1 is immunologically indistinguishable from LB_1 of the human LDLR (*9*). From this, we concluded that the two modules have the same disulfide bond connections. Here we have used the same approach to compare the conformation of the N-terminal module of rLB_{1-2} with the first ligand-binding module in the LDLR. The data in Figure 4 show that IgG-C7 binds to rLB_{1-2}, in the presence, but not in the absence of calcium. To confirm that the N-terminal module was recognized, LB_{1-2} was cleaved with pepsin and the N- and C-terminal fragments were purified by RP-FPLC. Dot blot analysis showed that IgG-C7 recognized LB_1(45) but not LB_2(40) (data not shown). These data show that LB_1 in rLB_{1-2} and in the LDLR are immunologically indistinguishable, suggesting that the structure and disulfide-bonding pattern for the N-terminal module in rLB_{1-2} is the same as that for LB_1 of the LDLR and rLB_1 (*9*).



Fragment	predicted M_r	experimental M_r
Cys(I)-Cys(III)	1871	1868
Cys(II)-Cys(V)	1332	1333
Cys(IV)-Cys(VI)	752	751
Cys(I)-Cys(III)-Cys(II)-Cys(V)	3068	3068
Cys(II)-Cys(V)-Cys(IV)-Cys(VI)	2011	2009

FIGURE 5: Analysis of elastase digests of $LB_2(40)$. The purified peptide fragment was digested with elastase as described in the text. The predicted cleavage sites are shown by dotted lines in the top panel. Disulfide bond connections are those predicted from rLB_2. Major cleavage fragments are labeled by the presence of cysteine residues. Arrows on either side of Gly58 indicate that this residue is part of the fragment designated by Cys(I)−Cys(III)−Cys(II)−Cys(V) in the table below. Cleavage products were separated by RP-FPLC and analyzed by mass spectrometry. Fragments with predicted and experimental molecular weights are shown in the table in the bottom panel.

Disulfide Bonds in LB_2. Since there are no conformation-specific monoclonal antibodies against LB_2, the arrangement of disulfide bonds in the C-terminal module of rLB_{1-2} was determined from an analysis of proteolytic fragments of RP-FPLC-purified peptic fragment $LB_2(40)$. $LB_2(40)$ contains six Gly−X peptide bonds, each of which is a potential cleavage site for elastase. Figure 5 illustrates the predicted fragments that would be formed if $LB_2(40)$ had the same disulfide bond connections as previously determined for rLB_2 (*6*). Cleavage of all of the Gly−X bonds (together with the pre-existing cleaved Phe67−Trp68 bond) is predicted to lead to three major fragments that contain the disulfide-bonded cysteine residues Cys(I)−Cys(III), Cys(II)−Cys(V), and Cys(IV)−Cys(VI), a short five-amino acid fragment (Ser80−Gly84) and glycine (Figure 5).

When $LB_2(40)$ was cleaved with elastase and the products analyzed by LC−MS, five major fragments were detected with molecular weights that ranged between 751 and 3068 (Figure 5). The molecular weights were consistent with those calculated for the three major cysteine-containing fragments, as well as for combinations of these fragments resulting from incomplete cleavage. The short fragment (Ser80−Gly84) was only detected as part of the much larger fragment Cys(II)−Cys(V)−Cys(IV)−Cys(VI). This combination of fragments cannot be derived from alternative disulfide-bonded structures. We conclude that the disulfide bond connectivity for the C-terminal ligand-binding module of rLB_{1-2} is the same as that previously reported for rLB_2 (*6*).

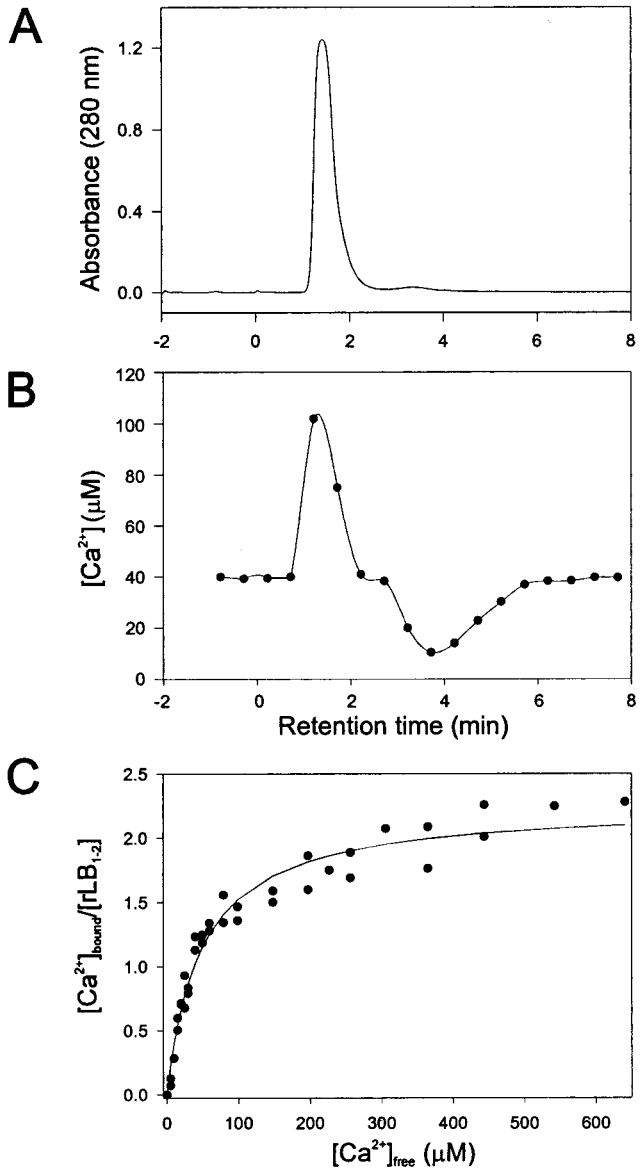


FIGURE 6: Analysis of calcium-binding by zonal chromatography. A PC3.2/10 fast desalting column was equilibrated with 50 mM Tris-HCl (pH 7.5), 0.15 M NaCl, 40 μM $CaCl_2$, and 0.5 nCi $^{45}CaCl_2$ per milliliter. A 90 μM solution of rLB_{1-2} (100 μL), in equilibration buffer, was applied to the column and eluted with at a flow rate of 200 μL/min. The elution profile was monitored spectrophotometrically at 280 nm. Fractions (100 μL) were collected and counted. (A) Elution profile for rLB_{1-2}. (B) Concentration of calcium in individual fractions. (C) Plot of $[Ca^{2+}]_{bound}/[rLB_{1-2}]$ vs the free calcium concentration. The data were fitted by least-squares analysis assuming identical independent sites ($n = 2.2$, $K_d = 48$ μM), as described in the text.

Calcium Binding. We have previously shown that the conformations of rLB_1 and rLB_2 are dependent on calcium, suggesting that these modules have a calcium-binding site (*8*, *23*). To determine the number of calcium-binding sites in rLB_{1-2}, and their affinity, we adapted the Hummel−Dreyer zonal chromatographic method (*14*, *24*) for use on a Smart System with a desalting column. The PC3.2/10 fast desalting column used for these measurements has a molecular weight cutoff of >5000; as a result, rLB_{1-2} elutes in the void volume, while calcium ions elute in the included column volume. Typical elution profiles for rLB_{1-2} and calcium are shown in Figure 6. The peptide eluted 1.15 min after injection and was quantified by its absorbance at 280 nm

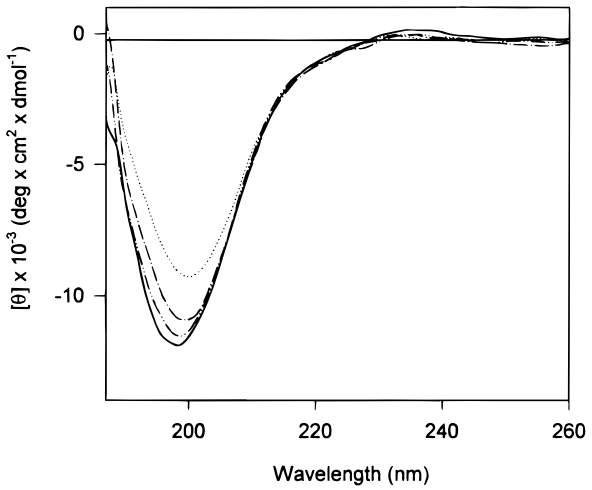


FIGURE 7: Effect of calcium on CD spectra of rLB_{1-2}. CD spectra were obtained in the absence of calcium ions (••) and in the presence of 40 μM (-•-), 400 μM (-••-), and 1 mM (−) calcium ions.

(panel A). Bound $^{45}Ca^{2+}$ coeluted with the rLB_{1-2} peak (panel B). This was followed by a trough that corresponds to a calcium-depleted zone. Because rLB_{1-2}, and hence its calcium complex, are excluded from the gel phase, they coelute. Consequently, the calcium deficit in this zone corresponds to the amount of calcium bound by rLB_{1-2}, and theoretically is equal to the excess calcium in the protein peak.

A plot of calcium bound per mole of rLB_{1-2} (ν) versus the total calcium concentration [Ca^{2+}] is shown in Figure 6C. These data were analyzed by nonlinear least-squares fitting (Table Curve 2D, Jandel Scientific) of the general binding function

$$\nu = \sum_{i=1}^{m} \frac{n_i[Ca^{2+}]/K_{di}}{1 - [Ca^{2+}]/K_{di}}$$

in which m different classes of sites (i) each bind n_i mol of calcium with a dissociation constant K_{di}. Assuming identical independent sites, this calculation predicts 2.2 ± 0.1 calcium-binding sites with a K_d of 48 ± 4 μM. This is consistent with two high-affinity calcium-binding sites, one for each ligand-binding module.

Circular Dichroism. The far-UV CD spectrum of rLB_{1-2} in the presence of 1 mM $CaCl_2$ (pH 7.5) displayed a strong negative band at 195 nm (Figure 7). In the absence of calcium (1 mM EDTA), the spectrum was qualitatively similar, but with a decrease in the intensity of the negative band. The differences between these spectra are most likely due to the effects of calcium on the structure of rLB_{1-2}. The CD profiles, and the changes induced by calcium, are similar to those reported for rLB_1 (*9*, *23*) and rLB_2 (*8*).

Analytical Ultracentrifugation. It is not known whether ligand-binding modules of the LDLR interact with each other to form a functional ligand-binding domain. However, the presence of exposed hydrophobic residues in the three-dimensional structures of rLB_1 and rLB_2 (*8*, *9*) has raised the possibility that such interactions may be a general property of ligand-binding modules. To address this question, we used analytical ultracentrifugation to measure the molecular weights of individual modules (rLB_1 and rLB_2) and of the concatemer. Intermolecular interactions are predicted to lead to deviations from calculated molecular weights. The results in Table 1 show that the experimental molecular weights of rLB_1, rLB_2, and rLB_{1-2}, in the presence or absence of calcium, were similar to those measured by electrospray mass spectrometry. We also measured the average molecular weight rLB_1 and rLB_2, in a 1:1 mixture (35 μM each) with and without calcium. Again, these were similar to those measured by electrospray mass spectrometry (results not shown). We conclude that LB_1 and LB_2 do not interact strongly under the experimental conditions used for ultracentrifugation and that the ability of ligand-binding modules to interact with each other is not a general property of these modules. We note that these data do not exclude specific interactions between neighboring or distant modules.

Table 1: Sedimentation Analyses[a]

			M_r	
peptide	[peptide] (mM)	[Ca^{2+}] (mM)	sedimentation	MS
rLB_{1-2}	0.019	0	10300 ± 900	9238 ± 1
rLB_{1-2}	0.018	10	9900 ± 800	
rLB_1	0.71	0	6100 ± 600	5220 ± 1
rLB_1	0.88	10	6300 ± 600	
rLB_2	0.42	0	4900 ± 500	4582 ± 1
rLB_2	0.42	10	5700 ± 500	

[a] Molecular weights were determined by analytical ultracentrifugation in the presence and absence of calcium ions. Molecular weights determined by electrospray mass spectrometry are shown for comparison.

NMR Analysis. The conformations of LB_1 and LB_2 in the concatemer and as isolated modules were compared by NMR spectroscopy. Sequence-specific assignments for rLB_{1-2} were made in the presence of 15 mM $CaCl_2$ (an ~4-fold molar excess) by reference to assignments for the individual modules (*8*, *9*) and were confirmed by the identification of sequential NOEs. Both calcium-binding sites of rLB_{1-2} ($K_d = 48 \pm 4$ μM) are saturated under these conditions. The chemical shifts of residues in rLB_{1-2} are not significantly different from those in the individual modules. This is readily seen by comparing the difference from random coil values of the α-proton chemical shifts, defined as the secondary chemical shift (*21*), of the individual modules and the concatemer (Figure 8). Reference chemical shifts for α-protons of the linker were those obtained for rLB_1 (*9*). The α-proton secondary chemical shift is a sensitive indicator of the secondary structure (*25*); thus, the preservation of these shifts in rLB_{1-2} is a reliable indicator that the conformations of the individual modules are maintained when ligated. The structure of the short linker region of four amino acids between the disulfide-bridged domains of the two modules was expected to be most sensitive to the presence of both modules. These residues were found to be poorly defined when expressed as a C-terminal extension of the first (*9*) or an N-terminal extension of the second module (*8*), and to have proton shifts similar to random coil values. In rLB_{1-2}, only relatively minor changes in the α-proton shifts are observed in these residues (Leu45, 0.11 ppm; Ser46, 0.10 ppm; Val47, 0.02 ppm; and Thr48, 0.03 ppm), suggesting that this linker retained most of its flexibility in the concatemer. The lack of significant changes in the α-proton chemical shifts, particularly in the linker region, suggests that the individual conformations of rLB_1 and rLB_2 are

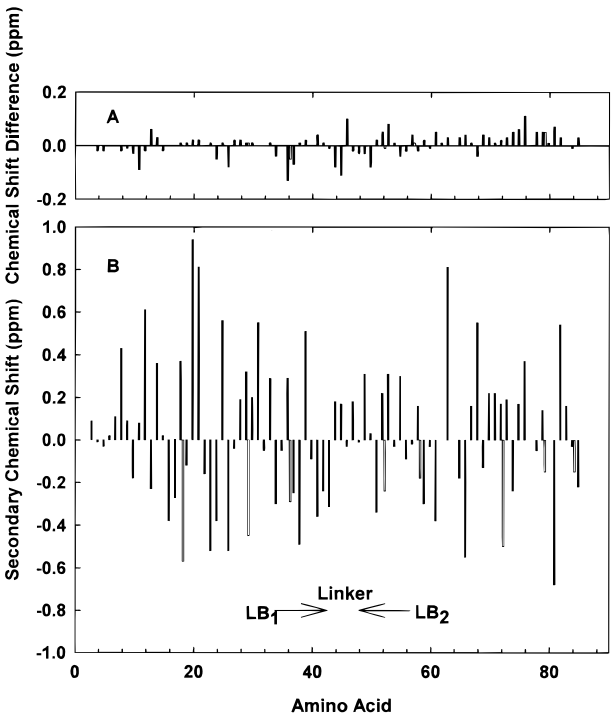


FIGURE 8: α-Proton chemical shifts of rLB_{1-2}, reported as the differences from random coil values (*22*) (panel B) and differences between corresponding chemical shifts for rLB_{1-2} and the individual modules rLB_1 and rLB_2 (panel A). Chemical shifts for protons in the linker region were those determined for rLB1. Shifts were measured from spectra recorded at 310 K and pH 6.1 for rLB_{1-2} and pH 6.3 for rLB_1 (*8*) and at 303 K and pH 5.5 for rLB_2 (*8*). The location of the linker that joins the last Cys residue of LB_1 with the first Cys residue of LB_2 is shown. Chemical shifts for pairs of Gly α-protons are shown as adjacent black and white bars.

maintained in the concatemer and there are few, if any, strong interactions between the two modules.

DISCUSSION

Members of the LDLR superfamily use discrete sets of cysteine-rich ligand-binding modules to bind multiple ligands. The LDLR uses its single string of seven modules to bind apoB-100, a large 500 000 Da protein associated with LDL, and apoE, a 34 000 Da protein associated with a number of different lipoproteins. Analyses of LDLRs with naturally occurring (*26*, *27*) and engineered (*12*, *13*) mutations have shown that the binding of apoB-100 requires a combination of modules 2−6, while the binding of apoE-containing lipoproteins depends critically on module 5. The LDLR is also remarkably selective. It distinguishes between naturally occurring isoforms of apoE (E2−4) that differ by single Cys → Arg substitutions. ApoE3 and apoE4 bind with high affinity, while apoE2 binds very weakly. An understanding of how seven ligand-binding modules bind apoB-100 and apoE, and how they distinguish between different isoforms of apoE, will require a description of the manner in which these modules interact with each other.

As a first step toward this goal, we have expressed and characterized a concatemer of the first two ligand-binding modules of the LDLR. This concatemer is not expected to bind lipoproteins since it lacks critical modules that are required for recognition of apoB-100 and apoE (see above). In the presence of calcium, the concatemer folded into a single, fully oxidized isomer. Several lines of evidence show that the disulfide bond connections are the same as those for the isolated modules (a 1−3, 2−5, and 4−6 pattern). First, proteolytic cleavage of the linker that joins LB_1 and LB_2 in the concatemer released N- and C-terminal modules, each with three disulfide bonds, demonstrating that all of the disulfide bonds were formed within the two modules. Second, monoclonal antibody IgG-C7 recognized the N-terminal module in the presence but not in the absence of calcium. This antibody only binds the calcium complex of the correct isomer of rLB_1. The structural basis for this specificity is twofold: (i) the correctly folded isomer of LB_1 only acquires an organized structure in the presence of calcium ions (*23*), and (ii) the coordination geometry of the calcium binding site is precisely controlled and is only present in the correct isomer of LB_1 (see below). Third, fragments formed by digestion of the C-terminal peptic fragment of rLB_{1-2} with elastase were those predicted from a 1−3, 2−5, and 4−6, but no other pattern of disulfide bonds.

The folding reaction that converts a complex mixture of isomers into the correct isomer is remarkably efficient, yielding >95% of the correct isomer. The efficiency with which longer strings fold into a single isomer is likely to depend on the properties of individual isomers. We have recently shown that LB_4 fails to fold into a single isomer in the presence of oxidized and reduced glutathione and calcium ions (unpublished data). This raises the possibility that some modules only fold efficiently in the presence of neighboring modules. In a recent independent study, Simmons et al. (*28*) reported the folding of a string of all seven modules of the ligand-binding domain into a form that bound LDL. Although the final yield was only 10%, the fact that this string recognized LDL well suggests that the majority of these modules exist as the correct isomer.

The folded concatemer binds two calcium ions with similar affinities ($K_d = 48 \pm 4$ μM), one for each module. The affinity for calcium is similar to that for the individual recombinant modules rLB_1 ($K_d = 10$ μM) and rLB2 ($K_d = 14$ μM), and like rLB_1 and rLB_2, the concatemer has a well-defined structure in the presence but not in the absence of calcium. The dual actions of calcium and three disulfide bonds produce a stable framework for ligand-binding modules that are likely to tolerate a range of environments within the cell, including the acidic pH of the endosome.

In the crystal structure of LB_5, calcium is coordinated with octahedral geometry by Asp25, Asp29, Asp35, and Glu36 and the backbone carbonyls of Trp22 and Gly27 (Figure 9) (*29*). Because the four acidic side chains are highly conserved in ligand-binding modules of the LDLR (Figure 9), the ligands for calcium in rLB_{1-2} are predicted to be the side chain carboxylates of Asp28, Glu32, Asp38, and Glu39 (for LB_1) and Asp71, Asp75, Asp81, and Glu82 (for LB_2). We have recently identified Asp28, Asp38, and Glu39 as the ligands for calcium in rLB_1, through the absence of chemical shift changes of their side chain protons between pH 3.9 and 6.8 (*23*). Surprisingly, the chemical shifts of the side chain protons of Glu32 did change with pH, indicating that its carboxylate group is not sequestered from the solvent through calcium coordination. We note that LB_1 is the only module in which the second position of the ligating amino acids is Glu instead of Asp, raising the possibility that steric effects due to the additional methylene group prevent its coordination to calcium. Since there are

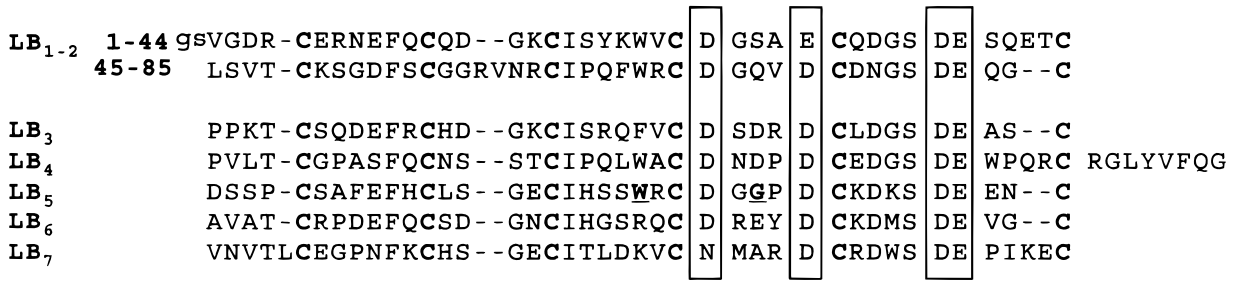


FIGURE 9: Sequence of rLB_{1-2} aligned with individual modules LB_3−LB_7. Two amino acids derived from the thrombin cleavage site of the GST fusion protein are lowercase. Conserved cysteine residues are bold. Acidic calcium-coordinating residues in LB_5, and corresponding residues in rLB_{1-2}, and other ligand-binding modules, are enclosed in boxes. Residues in LB_5 that coordinate calcium through backbone carbonyl oxygen atoms are bold and are underlined.

no nearby acidic residues in rLB_1 that can substitute for Glu32, a water molecule may be the sixth ligand (*23*). We have previously suggested that the loss of one of the peptide ligands may be the reason the affinity of calcium ions for rLB_1 is much lower than that for LB_5 ($K_d = -70$ nM). Although the ligands that coordinate calcium in LB_2 are not known, the fact that the affinities of calcium ions for LB_1 and LB_2 are similar suggests that both binding sites differ from that of LB_5. Finally, the similarity of the α-proton chemical shifts for the monomers and the concatemer strongly suggests that the same amino acids are used to coordinate calcium ions, and that the conformation of the calcium-binding loop is similar in the monomer and concatemer.

Like individual modules, the concatemer requires calcium ions to fold into a single isomer. The smaller divalent magnesium ions were unable to substitute for calcium. This was surprising since the LDLR has been reported to bind LDL in the presence of Mg^{2+} (*30*) and raises the possibility that magnesium ions are able to bind to one or more ligand-binding modules of a calcium-depleted LDLR. However, we have recently shown that Mg^{2+} does not bind to the correct isomer of rLB_1 (*23*). In view of the highly conserved coordinating residues within the calcium-binding loop of the LDLR and other members of the LDLR family, we surmise that the coordination geometry defined in the crystal structure of LB_5 will be conserved in other ligand-binding modules. This makes it unlikely that magnesium ions bind to any of the ligand-binding modules. Finally, we note that the requirement for calcium is not restricted to in vitro folding. Cell culture experiments have shown that ER-located LDLRs that have been reduced with DTT can only fold into a fully functional receptor in the presence of calcium ions (A. Ozinsky and D. R. van der Westhuyzen, personal communication).

The conformations of autonomously folding domains from modular proteins such as the LDLR are generally considered to be independent of the remainder of the protein. However, specific interactions between domains often exist and can lead to structural changes that are critical to the function of a protein. The close similarity between the α-proton chemical shifts for rLB_1, rLB_2, and the concatemer suggests that the conformations of LB_1 and LB_2 in these molecules are very similar and that few, if any, strong interactions exist between the two modules. The absence of strong interactions is further supported by ultracentrifugation data that showed no evidence for intermolecular interactions between rLB_1, rLB_2, and rLB_{1-2} molecules. The small differences between the α-proton chemical shifts for the concatemer and the monomers, and their similarity to random coil values, strongly suggest that the linker in the concatemer is flexible. The modules of rLB_{1-2} therefore appear to be structurally independent and to be joined by a flexible linker. The folded form of rLB_{1-2} provides a model for more detailed studies of the effects of mutations on the folding of ligand-binding modules.

The degree to which modules in other proteins interact is highly variable. The fibronectin type I (F1) and epidermal growth factor (EGF)-like modules of the tissue-type plasminogen activator interact intimately in the F1−EGF pair, with little or no reorientation of the two modules with respect to one another (*21*). In contrast, the EGF and kringle modules from the N-terminal fragment of the urokinase-type plasminogen activator are structurally independent domains that rotate independently of each other (*31*). The two domains of the human complement factor H (H15 and H16) provide an intermediate example, where a limited number of interactions are observed, but which are insufficient to totally define the relative orientations of the two domains (*32*). The modules of rLB_{1-2} thus resemble the EGF and kringle modules of the urokinase-type plasminogen activator; they are joined by a flexible domain, are structurally independent, and do not interact strongly with each other.

What do the current data tell us about the structure of the ligand-binding domain of the LDLR? If all of the modules exist as structurally independent domains that are linked by a flexible linker, the ligand-binding domain will resemble beads on a string. Such a structure would provide a great deal of flexibility and would allow the modules to rearrange themselves to bind to a variety of ligands, including apoB-100, apoE, and lipoprotein lipase.

REFERENCES

1. Goldstein, J. L., Brown, M. S., Anderson, R. G., Russell, D. W., and Schneider, W. J. (1985) *Annu. Rev. Cell Biol. 1*, 1.
2. Brown, M. S., and Goldstein, J. L. (1986) *Science 232*, 34.
3. Goldstein, J. L., Hobbs, H. H., and Brown, M. S. (1995) in *The Metabolic and Molecular Basis of Inherited Disease* (Scriver, C. S., Beaudet, A. L., Sly, W. S., and Valle, D., Eds.) pp 1981−2030, McGraw-Hill, New York.
4. Sudhof, T. C., Goldstein, J. L., Brown, M. S., and Russell, D. W. (1985) *Science 228*, 815.
5. Bieri, S., Djordjevic, J. T., Daly, N. L., Smith, R., and Kroon, P. A. (1995) *Biochemistry 34*, 13059.
6. Bieri, S., Djordjevic, J. T., Jamshidi, N., Smith, R., and Kroon, P. A. (1995) *FEBS Lett. 371*, 341.

7. Mahley, R. W. (1988) *Science 240*, 622.
8. Daly, N. L., Djordjevic, J. T., Kroon, P. A., and Smith, R. (1995) *Biochemistry 34*, 14474.
9. Daly, N. L., Scanlon, M. J., Djordjevic, J. T., Kroon, P. A., and Smith, R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6334.
10. Djordjevic, J. T., Bieri, S., Smith, R., and Kroon, P. A. (1996) *Eur. J. Biochem. 239*, 214.
11. Blacklow, S. C., and Kim, P. S. (1996) *Nat. Struct. Biol. 3*, 758.
12. Esser, V., Limbird, L. E., Brown, M. S., Goldstein, J. L., and Russell, D. W. (1988) *J. Biol. Chem. 263*, 13282.
13. Russell, D. W., Brown, M. S., and Goldstein, J. L. (1989) *J. Biol. Chem. 264*, 21682.
14. Hummel, J. P., and Dreyer, W. J. (1962) *Biochim. Biophys. Acta 63*, 530.
15. Cassasa, E. F., and Eisenberg, H. (1964) *Adv. Protein Chem. 19*, 287.
16. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, S. J., and Horton, J. C., Eds.) pp 90−125, Royal Society of Chemistry, Cambridge, England.
17. Cohn, E. J., and Edsall, J. T. (1943) in *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions* (Cohn, E. J., and Edsall, J. T., Eds.) pp 370−381, Reinhold, New York.
18. Bax, A., and Davis, D. G. (1997) *J. Magn. Reson. 65*, 355.
19. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546.
20. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661.
21. Smith, B. O., Downing, A. K., Dudgeon, T. J., Cunningham, M., Driscoll, P. C., and Campbell, I. D. (1994) *Biochemistry 33*, 2422.
22. Merutka, G., Dyson, H. J., and Wright, P. E. (1995) *J. Biomol. NMR 5*, 14.
23. Atkins, A. R., Brereton, I. M., Kroon, P. A., Lee, H. T., and Smith, R. (1998) *Biochemistry 37*, 1622.
24. Cann, J. R., Rao, A. G., and Winzor, D. J. (1989) *Arch. Biochem. Biophys. 270*, 173.
25. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647.
26. Hobbs, H. H., Russell, D. W., Brown, M. S., and Goldstein, J. L. (1990) *Annu. Rev. Genet. 24*, 133.
27. Sass, C., Giroux, L., Lussier-Cacan, S., Davignon, J., and Minnich, A. (1995) *J. Biol. Chem. 264*, 25166.
28. Simmons, T., Newhouse, Y. M., Arnold, K. S., Innerarity, T. L., and Weisgraber, K. H. (1974) *J. Biol. Chem. 272*, 25531.
29. Fass, D., Blacklow, S., Kim, P. S., and Berger, J. M. (1997) *Nature 388*, 691.
30. Goldstein, J. L., and Brown, M. S. (1977) *Annu. Rev. Biochem. 46*, 897.
31. Hansen, A. P., Petros, A. M., Meadows, R. P., Netteshiem, D. G., Mazar, A. P., Olejniczak, E. T., Xu, R. X., Pederson, T. M., Henkin, J., and Fesik, S. W. (1994) *Biochemistry 33*, 4847.
32. Barlow, P. N., Steinkasserer, A., Norman, D. G., Kieffer, B., Wiles, A. P., Sim, R. B., and Campbell, I. D. (1993) *J. Biomol. NMR 232*, 268.

BI980452C