

# Folding and Binding Integrity of Variants of a Prototype Ligand-Binding Module from the LDL Receptor Possessing Multiple Alanine Substitutions<sup>†</sup>

Dunia Abdul-Aziz, Carl Fisher, Natalia Beglova, and Stephen C. Blacklow\*

Department of Pathology, Harvard Medical School and Brigham & Women's Hospital,  
77 Avenue Louis Pasteur, Boston, Massachusetts 02115

Received November 17, 2004; Revised Manuscript Received February 3, 2005

**ABSTRACT:** The LA repeats that comprise the ligand-binding domain of the LDL receptor are among the most common autonomously structured extracellular modules found in the nonredundant protein sequence database. Here, we investigate the information content of the amino acid sequence of a typical LA module by constructing sequences with alanine residues at nonconserved positions in the module. Starting with the sequence of the fifth ligand-binding repeat of the LDL receptor (LA5), we created generic LA modules with alanine substitutions of nonconserved residues in only the N-terminal lobe, only the C-terminal lobe, and throughout both lobes of the module. LA variants with alanine residues at as many as 18 of 37 positions fold to a preferred disulfide isomer in the presence of calcium. Indeed, the six cysteines, the C-terminal calcium coordinating residues, two hydrophobic residues involved in packing, two glycines, and five other residues that form side chain–intramodule hydrogen bonds are alone sufficient to specify the fold of an LA module when alanine residues are present at all other positions. The LA variants with multiple alanines in either the N- or C-terminal lobe were then exploited to identify residues of LA5 that contribute to the binding of apoE-containing ligands in LDL receptor-derived “minireceptors”, implicating nonconserved residues of the N-terminal lobe of LA5 in recognition of apoE–DMPC. Our library of LA modules with multiple alanine substitutions should be generally useful for probing the roles of nonconserved side chains in ligand recognition by proteins of the LDL receptor family.

The LA repeats that comprise the ligand-binding domain of the LDL<sup>1</sup> receptor are among the most common autonomously structured extracellular modules found in the nonredundant protein sequence database (1). Though first identified in the LDL receptor, receptors that contain LA modules are now known to participate in a diverse array of biological processes ranging from viral infection to Wnt signal transduction and to brain development (2).

Because LA modules are so ubiquitous, we sought to understand fundamental interrelationships among their sequence, structure, and function. In the work reported here, we first identify side chains that participate in specifying the native LA module fold by examining the folding behavior of designed modules in which the information content of a prototype module, the fifth ligand-binding repeat of the LDL receptor (LA5), is systematically reduced by alanine-saturation mutagenesis at nonconserved positions (3–5). Using the binding of apoE–DMPC by LDL receptor-derived minire-

ceptors as a model system (6), we then substitute folding-competent mutants in place of LA5 in domain swap experiments to evaluate the functional consequences of reducing the information content at those nonconserved residues tolerant of alanine substitution.

Structures of various LA modules, determined by both NMR spectroscopy and X-ray crystallography, are largely devoid of secondary structure elements, but all exhibit a well-defined tertiary fold specified primarily by three disulfide bonds and the coordination of a calcium ion (7–16). In the structure of LA5 (Figure 1A), a prototypical module, the six conserved cysteine residues form disulfide bonds organized in a I–III (Cys 176–Cys 188), II–V (Cys 183–Cys 201), and IV–VI (Cys 195–Cys 210) arrangement and form two loops which define the N-terminal and C-terminal “lobes” of the module (9). The N-terminal lobe consists of residues 175–192 and contains the disulfide bond linking the first and third cysteine residues as well as a very short antiparallel  $\beta$ -sheet. Within the C-terminal lobe (residues 193–211), the four carboxylate groups on the side chains of conserved acidic residues (Asp 196, Asp 200, Asp 206, and Glu 207) and two carbonyl oxygens (of Trp 193 and Gly 198) in the protein backbone are arranged in an octahedral geometry to coordinate the calcium ion.

Though structures have been solved for several LA modules, less is known about how the primary sequence of an LA module specifies its native fold. These studies focus on the design of an LA module with the minimum amount of side chain information required to specify the native fold. We hypothesized that (i) the most highly conserved residues

<sup>†</sup> Supported by NIH Grant No. R01 HL-61001. S.C.B. is a Pew Scholar in the Biomedical Sciences and an Established Investigator of the American Heart Association.

\* To whom correspondence should be addressed. Phone, 1-617-525-4413; fax, 1-617-525-4414; e-mail, sblacklow@rics.bwh.harvard.edu.

<sup>1</sup> Abbreviations: apoE, apolipoprotein E; BPTI, bovine pancreatic trypsin inhibitor;  $\beta$ -VLDL,  $\beta$ -migrating very low-density lipoprotein; DMPC, dimyristoyl phosphatidylcholine; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HPLC, high-pressure liquid chromatography; LDL, low-density lipoprotein; LDLR, low-density lipoprotein receptor; LA, LDL receptor type-A module; LA5, ligand-binding module 5 of the low-density lipoprotein receptor; MALDI–TOF, matrix-assisted laser desorption/ionization–time of flight; TFA, trifluoroacetic acid.

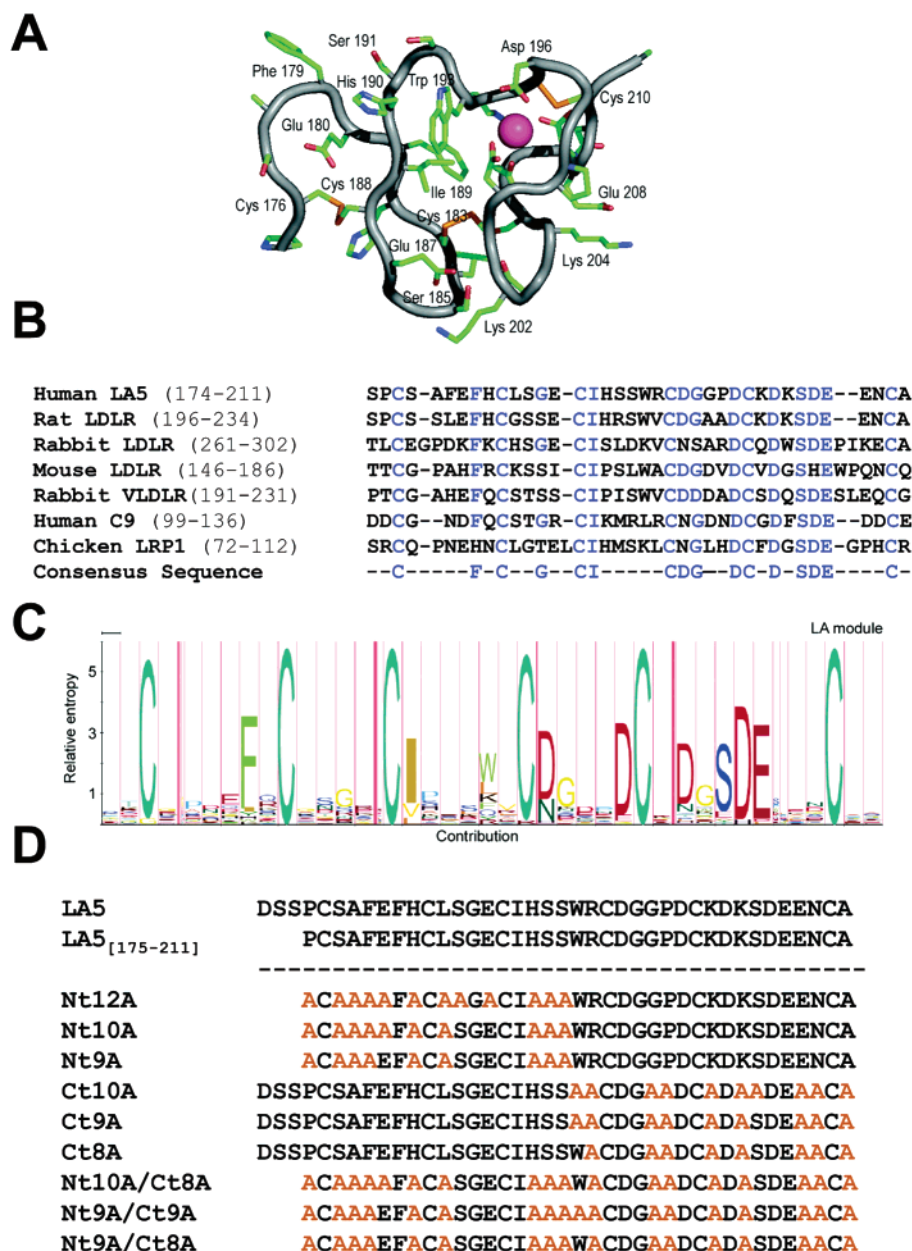


FIGURE 1: (A) Ribbon diagram of the LA5 crystal structure (PDB entry 1AJJ). The structure shows the arrangement of the three disulfide bonds, the calcium coordination site, and important hydrophobic contacts. (B) Sample of the amino acid sequences taken from the seed alignment of related LA modules (PFAM database) comparing human, rat, rabbit, and mouse LDL receptor, rabbit VLDL receptor, the C9 component of complement, and chicken LDL receptor-related protein 1 (LRP1). The consensus sequence is shown at the bottom, based on residues conserved in greater than 60% of the interspecies LA modules analyzed. (C) Hidden Markov model logo of the LA module family schematically illustrating the degree of information content at each position. The frequency of a residue at a particular position is proportional to its height in the model (23). (D) Amino acid sequences of the wild-type and mutant LA5 modules purified for these studies. The LA5 variants contain alanine substitutions for nonconserved residues in the N-terminal lobe (Nt12A, Nt10A, Nt9A), in the C-terminal lobe (Ct10A, Ct9A, Ct8A), and throughout both lobes of the module (Nt10A/Ct8A, Nt9A/Ct9A, Nt9A/Ct8A). Alanine residues are highlighted in orange.

among LA modules are required to specify the fold and that (ii) the most highly variable positions, identified in a sequence alignment of multiple LA modules, do not contribute to proper folding and would thus tolerate substitution with the generic amino acid alanine.

To test this hypothesis, nonconserved residues that exhibit the highest sequence entropy among aligned LA modules were replaced with alanines. We find that LA modules with as many as 18 alanines exhibit calcium-dependent folding to a single disulfide-bonded isomer and bind calcium with nativelike affinity. When used to probe the contribution of nonconserved residues of LA5 to ligand binding, domain

swapping of LA5 with an N-terminal alanine-rich variant prevents binding, while a domain swap with a C-terminal alanine-rich module does not disrupt binding. Together, these results provide important insights into the specific determinants of both the folding and binding properties of LA modules.

## EXPERIMENTAL PROCEDURES

**Protein Expression and Purification.** The pMM-LA5 vector contains the LA5 gene fragment of the LDLR (encoding residues 172–211) flanked by *Hind*III and *Bam*HI restriction sites and a TrpLE sequence that directs the fusion

protein to inclusion bodies (17). The N-terminal Asp-Ser-Ser sequence was deleted from pMM-LA5 by PCR mutagenesis to make the LA5<sub>[175–211]</sub> module used in these studies (encoding residues 175–211), since these three residues are disordered in the LA5 crystal structure and do not influence folding. A series of alanine-substituted variants of either LA5 or LA5<sub>[175–211]</sub> (see Figure 1D) were then constructed using a synthetic gene approach and inserted into the expression vector in place of the LA5 sequence on a *Hind*III–*Bam*HI fragment. Each construct was named according to whether alanines were introduced into the N-terminal lobe (Nt), C-terminal lobe (Ct), or both lobes (Nt/Ct). Nt12A, Nt10A, and Nt9A have 12, 10, and 9 alanine residues, respectively, in the N-terminal lobe of LA5<sub>[175–211]</sub>; Ct10A, Ct9A, and Ct8A have 10, 9, and 8 alanines, respectively, in the C-terminal lobe of LA5. Nt10A/Ct8A, Nt9A/Ct9A, and Nt9A/Ct8A represent chimeras with 17 or 18 alanines substituted in both lobes of LA5<sub>[175–211]</sub> as indicated. The identity of each DNA construct was confirmed by DNA sequencing.

LA5, LA5<sub>[175–211]</sub>, and their variants were expressed in *Escherichia coli* strain BL21(DE3) as a (His)<sub>9</sub>–{TrpLE}–Met–{LA} fusion protein, with an N-terminal (His)<sub>9</sub> tag and the methionine residue separating the TrpLE sequence from the LA module. Each fusion protein was recovered from inclusion bodies after sonication in detergent as described (17). Briefly, inclusion bodies were dissolved overnight in 10 mM Tris buffer (pH 8.0) containing 6 M guanidine hydrochloride and 10 mM oxidized 2-mercaptoethanol, and the fusion protein was precipitated by 10-fold dilution with water. The precipitated fusion protein was resuspended in formic acid and cleaved with cyanogen bromide to release the LA fragment. This cleaved product was dialyzed exhaustively against 5% of acetic acid and lyophilized. After the lyophilized sample was redissolved in 10 mM Tris buffer (pH 8.0) containing 6 M guanidine hydrochloride and 5 mM DTT, the N-terminal tag was removed from solution by batch-binding of the sample to Ni<sup>2+</sup>–NTA agarose (Qiagen). For most preparations, the purified protein was refolded under conditions permitting disulfide exchange before final purification by HPLC. Exhaustive dialysis was performed at 4 °C for 72 h against 50 mM Tris buffer, pH 8, containing 10 mM CaCl<sub>2</sub>, 2 mM L-cysteine, and 0.5 mM L-cystine (buffer A), with daily buffer changes (6). Each protein was then purified to homogeneity by reversed-phase HPLC on a Vydac C-18 column using a linear water–acetonitrile gradient established by two solvents: water and 0.1% of TFA (solvent A) and 90% of acetonitrile, 10% of water, and 0.1% of TFA (solvent B). The masses of all purified proteins were confirmed by MALDI–TOF mass spectrometry, and each protein was stored in lyophilized form at –80 °C.

For constructs containing the native tryptophan residue, protein concentration was measured by protein absorbance (calculated as the difference between absorbance at 280 and 320 nm) in 20 mM sodium phosphate buffer, pH 6.5, containing 6 M guanidine hydrochloride. Molar extinction coefficients were calculated using the Protparam program based on tryptophan absorbance (18, 19). For the other proteins, concentration was estimated using the Bradford protein assay in which the absorbance measurement (at 595 nm) of the protein in detergent was correlated with concentration according to a set of BSA standards (BioRad).

Expression constructs for synthesis of an LA3–6 “mini-receptor” with or without a C-terminal myc epitope tag have been described previously (6). These constructs express the LA3–6 minireceptor as a glutathione S-transferase (GST) fusion, with a tobacco etch virus (TEV) protease cleavage site (E–N–L–T–F–Q–G) between GST and the LA3–6 sequence. Substitution of the Nt9A and Ct8A sequences in place of the LA5 module of these minireceptors was achieved by two cycles of PCR-directed mutagenesis. The identities of the resulting constructs were confirmed by DNA sequencing.

The native and mutant minireceptors were expressed in *E. coli* BL21(DE3) cells during log phase growth by induction with IPTG. Cells were harvested by centrifugation and lysed by sonication as described (6). The soluble fusion proteins were captured on Glutathione Sepharose beads (Pharmacia) and cleaved with TEV protease in Tris-buffered saline (TBS) supplemented with 0.5 mM EDTA and 1 mM DTT to release the desired minireceptor from the beads. The supernatant was separated from the beads by centrifugation, and TEV protease was then removed from solution by capture onto Ni<sup>2+</sup>–NTA agarose beads (Qiagen) via its hexahistidine tag followed by removal of the beads by filtration through fritted glass. Each minireceptor was then refolded under conditions permitting disulfide exchange by exhaustive dialysis against buffer A. The extent of disulfide bond formation was monitored by analytical reversed-phase HPLC using a Vydac C-18 column. For each minireceptor, the predominant disulfide isomer was purified to apparent homogeneity by reversed-phase HPLC using a shallow acetonitrile gradient, and protein concentration measurements based on tryptophan and tyrosine absorbance were determined as above.

N–apoE (residues 1–191 of human apolipoprotein E with an N-terminal hexahistidine tag) was purified as described previously (6). Discoidal complexes of N–apoE were prepared by bath sonication with dimyristoylphosphatidylcholine (DMPC) as described (6). Some disc preparations were further purified away from uncomplexed N–apoE by gel filtration over a Superdex 300 column (Pharmacia) and were then concentrated by ultrafiltration (Vivascience). An amount of 95% or more of the N–apoE was incorporated into discs as judged by native gel electrophoresis (Phast gel, Pharmacia), which was also used to estimate the size of discoidal complexes by comparison with molecular weight standards.

**Disulfide Exchange Experiments.** The calcium dependence of folding was investigated for each LA module by incubation under conditions that allow for disulfide exchange. Purified wild-type or mutant LA5 proteins (20 μM) were equilibrated under anaerobic conditions at room temperature in a redox buffer consisting of 50 mM Tris (pH 8.5), reduced glutathione (500 μM), oxidized glutathione (250 μM), and either calcium chloride (1 mM), EDTA (100 μM), or guanidine hydrochloride (6 M). After 48 h, the samples were acidified by addition of TFA (to a final concentration of 0.1%) to arrest disulfide exchange (except for the Nt12A variant which was acidified by addition of acetic acid to 5% v/v). The distribution of disulfide isomers was analyzed by reversed-phase HPLC on an analytical Vydac C-18 column using a linear acetonitrile gradient (0.1% of solvent B per minute) starting from approximately 23% of solvent B. The



distribution of isomers was indistinguishable for LA5<sub>[175–211]</sub>, Nt9A, or Ct8A, regardless of whether the rearrangements were initiated with reduced or oxidized protein, indicating that the isomer distribution was likely to be at equilibrium for each protein tested.

**NMR Spectroscopy.** <sup>15</sup>N-Labeled LA5<sub>[175–211]</sub>, Nt10A, Ct9A, or Nt9A/Ct9A proteins were prepared by inducing expression in *E. coli* grown in M9 minimal medium supplemented with <sup>15</sup>N ammonium chloride and purified as above. Each NMR sample contained about 0.7 mM protein in 90% H<sub>2</sub>O/10% D<sub>2</sub>O, 2.5 mM bis-Tris buffer, pH 6.9, and 10 mM CaCl<sub>2</sub>. This calcium ion concentration, which exceeds the calcium dissociation constant by several orders of magnitude, is sufficient to ensure saturation of all Ca<sup>2+</sup>-binding sites. <sup>15</sup>N-HSQC NMR spectra were acquired at 298 K on a Bruker 600 MHz spectrometer equipped with a triple resonance cryoprobe. Spectra were processed and analyzed with GIFA v4.3 (20). Resonance assignments for LA5<sub>[175–211]</sub> were transferred from prior assignments for the LA5–LA6 pair by analyzing HNCA, HNCOCa, and HNCACB data (21).

**Calcium Affinity Measurements.** For LA5<sub>[175–211]</sub>, Nt10A, Nt9A, Ct8A, and Nt9A/Ct8A, the calcium affinity of the folded module was measured by isothermal titration calorimetry (ITC). Calcium was removed from the folded LA modules prior to titration by reversed-phase HPLC purification under acidic conditions (the affinity of the LA modules for calcium is negligible at pH 2). Calcium titrations were performed at pH 5.6 in Chelex-treated MES buffer at 25 °C using a MicroCal VP-ITC calorimeter in a 1.3 mL reaction vessel. A CaCl<sub>2</sub> stock solution (0.5–2 mM) was added in 3–5 μL increments to various refolded LA5 variants present at concentrations from 80 to 120 μM. Calorimetry data were analyzed with the program Origin 5.0 (OriginLab Corp.).

**Co-Immunoprecipitation Binding Assay.** Co-immunoprecipitation of complexes between different LA3–6 minireceptors (wild-type LA3–6, LA3–6 with LA5 replaced by Nt9A, or LA3–6 with LA5 replaced by Ct8A) and apoE–DMPC was performed as previously described (6). In competition binding experiments, immunoprecipitations were carried out with the myc-tagged, wild-type LA3–6 minireceptor in the presence of 1, 5, or 20 molar equiv of untagged competitor (either wild-type LA3–6, LA3–6 with LA5 replaced by Nt9A, or LA3–6 with LA5 replaced by Ct8A).

## RESULTS

**Design of LA5 Variants.** For the purpose of designing alanine-saturated variants of LA5, we used the seed alignment of 53 LA modules in the PFAM database (Figure 1B,C). We considered a residue to be conserved if it appeared in more than 60% of the sequences at a given position. According to this definition, the conserved residues (Figure 1B, bottom line) include the six cysteines essential for defining the fold of the module (22), four acidic residues (Asp 196, Asp 200, Asp 206, and Glu 207) that coordinate a calcium ion needed to maintain the structural integrity of the module (9, 17), two hydrophobic residues (Phe 181 and Ile 189) implicated in packing of the N-terminal lobe, two glycine residues (Gly 186 and Gly 197), and two non-calcium coordinating polar residues in the C-terminal lobe (Asp 203 and Ser 205). A sequence logo plot (23) of the alignment highlights the residues that are most highly conserved (Figure 1C).

We divided the LA module into N- and C-terminal lobes, initially restricting alanine substitutions to only one lobe or the other. Then, the alanine substitutions from folding-competent sequences in each class of mutants (Nt or Ct) were combined to create sequences with alanine residues distributed throughout the whole module (Figure 1D).

To determine whether the different LA modules fold properly, we examined whether calcium was required to guide formation of a preferred disulfide isomer. We took this approach to assess the folding competence of our designed LA modules because the calcium dependence of folding is a hallmark of LA modules: wild-type LA modules (like LA5) form native disulfide bonds when calcium is present in the redox buffer, but a distribution of isomers is formed when calcium is removed from the folding buffer by addition of EDTA or when folding is carried out in denaturant (13, 15, 17, 24–26). As expected, the identical calcium dependence of folding is also observed for LA5<sub>[175–211]</sub> (Figure 2A).

In the first variant module tested, all 12 nonconserved residues in the N-terminal lobe of LA5<sub>[175–211]</sub> were replaced by alanines to create the Nt12A protein (Figure 1D). In contrast to wild-type LA5<sub>[175–211]</sub>, however, the Nt12A polypeptide did not fold to a preferred disulfide isomer under conditions permitting disulfide rearrangement, regardless of whether calcium was present (Figure 2B). The distribution of Nt12A isomers at equilibrium is similar in the presence or absence of calcium, as judged by the similarity in the HPLC chromatograms after folding in calcium (50 mM) or EDTA (0.1 mM; Figure 2B, top two traces). Refolding of an equivalent quantity of protein in a redox buffer containing both calcium ions in molar excess (50 mM) and denaturant (4 M guanidine hydrochloride) also yields a similar distribution of multiple isomers (bottom trace), but more protein is recovered, suggesting that Nt12A aggregates when the denaturant is not present.

Because Nt12A did not fold to a unique disulfide isomer in the presence of calcium, we designed two more conservative variants, Nt10A and Nt9A, with 10 and 9 alanines in the N-terminal lobe, respectively. The Nt10A mutant retains both Ser 185 and Glu 187 from the native LA5 sequence because the X-ray crystal structure (9) identifies a hydrogen bond between Glu 187 and the hydroxyl side chain group of Ser 185 (Figure 1D) and because an LA5 module bearing the Glu 187 to Lys mutation found in familial hypercholesterolemia patients fails to fold to a unique disulfide isomer in the presence of calcium *in vitro* (17). Finally, the Nt9A construct also retains the glutamic acid residue at position 180 because it exhibits 42% conservation among the aligned LA modules.

Remarkably, both Nt10A and Nt9A exhibit calcium dependent folding under conditions permitting disulfide exchange, with the preferred isomer exhibiting an HPLC retention time similar to that of LA5<sub>[175–211]</sub> (Figure 2C,D). The existence of a small nonnative isomer peak for Nt10A when refolded in 1 mM calcium buffer reveals that calcium concentrations higher than 1 mM are required to drive folding to completion (Figure 2C, top trace). In comparison, the Nt9A mutant exhibits complete folding to a single isomer at 1 mM calcium concentration (Figure 2D, top trace).

Next, we created C-terminal alanine substitutions using similar criteria. The Ct9A construct has all nine nonconserved

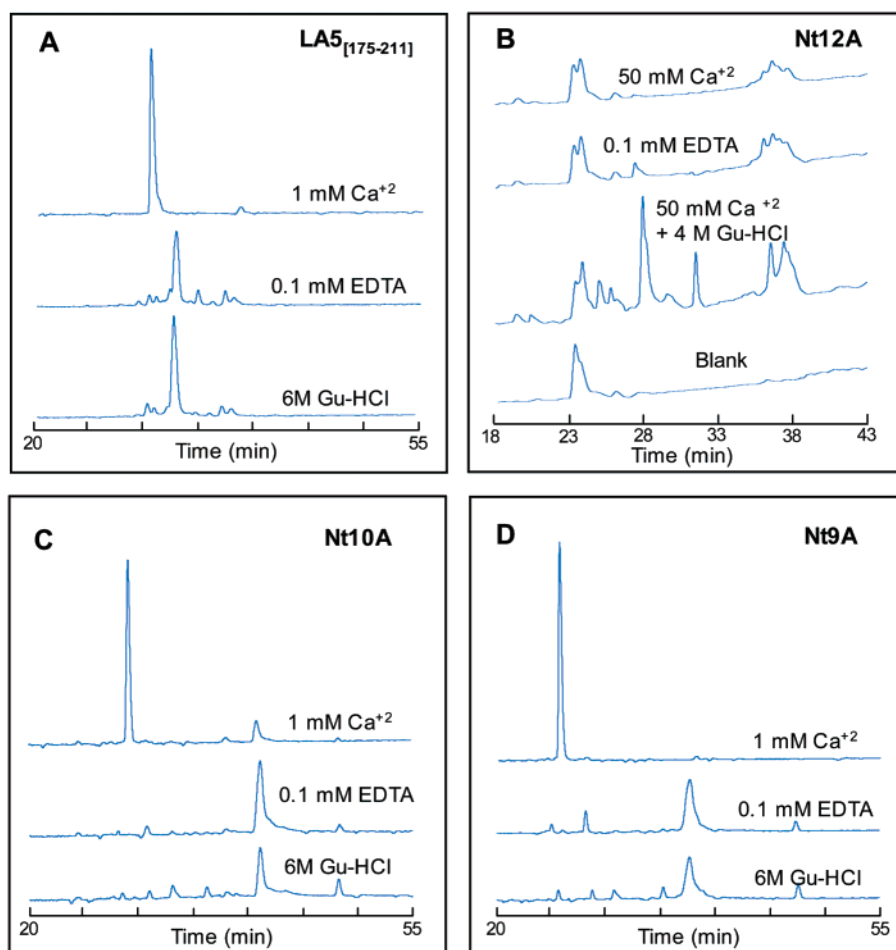


FIGURE 2: HPLC chromatograms of products formed after oxidative folding of LA5<sub>[175-211]</sub> (A), Nt12A (B), Nt10A (C), and Nt9A (D). The proteins were incubated in a redox buffer containing calcium, EDTA, or denaturant as described in Experimental Procedures.

residues replaced by alanines. Ct8A has eight alanine residues, retaining Trp 193 because it is conserved in 47% of the aligned sequences and because its backbone oxygen atom participates in calcium coordination. Finally, Ct10A has a 10th alanine residue in place of Ser 205 in addition to the nine alanines from Ct9A, probing its role in contributing to proper folding. The reason for the conservation of Ser 205 is not immediately apparent, though its side chain appears to form two hydrogen bonds, as a hydrogen bond acceptor from the backbone amide of residue 184 and as a hydrogen bond donor to the side chain of Asp 203 (9).

Each of the three C-terminal alanine variants was purified to homogeneity and examined to determine whether they were capable of forming a preferred disulfide isomer in the presence of calcium. Both Ct8A and Ct9A fold to a preferred disulfide isomer in the presence of 1 mM calcium (Figure 3A,B), the concentration typically used to screen for the calcium-dependence of folding of the different alanine-substituted variants. In contrast, the isomer distribution of Ct10A does not strongly favor a single isomer until 10 mM calcium is present in the folding buffer (Figure 3C). When incubated in a redox buffer containing either EDTA or denaturant, the bulk of the Ct8A protein (as well as the Nt10A/Ct8A, Nt9A/Ct9A, and Nt9A/Ct8A proteins; see below) does not elute from the HPLC column, and a broad distribution of small-intensity peaks is seen instead. This observation suggests that most of the Ct8A forms disulfide-bonded aggregates when folding is

attempted in the absence of calcium (or is perhaps proteolyzed by trace amounts of proteases active under these folding conditions).

Finally, mutants possessing substitutions throughout the entire module were constructed based on the folding ability of the single-lobe-substituted mutants above. To create LA modules with alanines distributed throughout both lobes of the module, N-terminal and C-terminal mutants folding to a predominant disulfide isomer were combined to create the Nt10A/Ct8A, Nt9A/Ct9A, and Nt9A/Ct8A modules. Each of these variants, which contain either 17 or 18 alanine residues, forms a single, homogeneous species when folded in the presence of 1 mM calcium (Figure 4A–C, top traces), whereas a minimal amount of protein is recovered when folding is attempted in EDTA or denaturant (Figure 4A–C, middle and bottom traces). The difference in elution times among the folded polypeptides is likely a consequence of their varying hydrophobicities, which is influenced in a large part by the presence or absence of Trp 193.

**NMR Spectroscopy of Native and Mutant LA5 Modules.** To determine unambiguously whether the predominant peak in the HPLC chromatogram indeed corresponds to a single-folded species (i.e., a unique, well-folded disulfide isomer), we analyzed mutant LA5 modules by NMR spectroscopy. Each of three mutants tested, Nt10A, Ct9A, and Nt9A/Ct9A, exhibits an HSQC spectrum as well-dispersed as that of native LA5<sub>[175-211]</sub>, indicating that the mutants are as well-folded as the wild-type protein (Figure 5A–C). The number

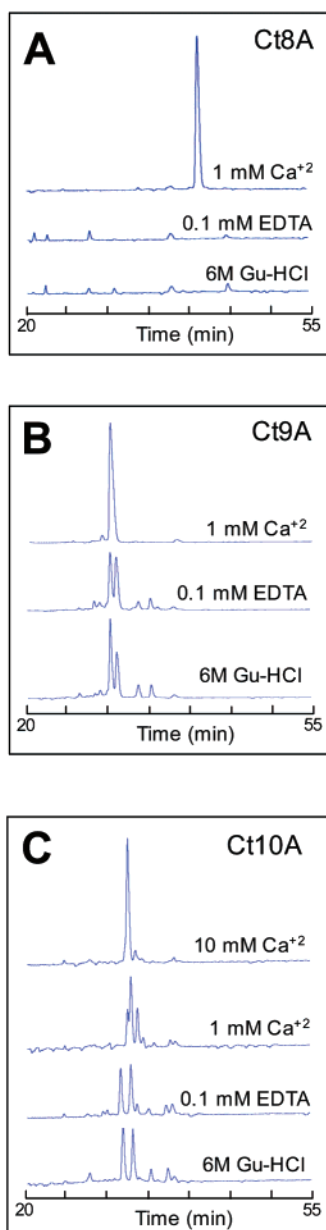


FIGURE 3: HPLC chromatograms of products formed after oxidative folding of Ct8A (A), Ct9A (B), and Ct10A (C). The proteins were incubated in a redox buffer containing calcium, EDTA, or denaturant as described in Experimental Procedures.

of peaks observed for each mutant corresponds to that expected for a single homogeneous species with no evidence for molecular heterogeneity.

To evaluate the effect of the alanine substitutions more precisely, the backbone amide peak of each unmutated residue in the spectrum of the mutant was mapped to the closest peak in the spectrum of the wild-type protein, and the change in chemical shift was calculated according to eq 1:

$$\{(\Delta\delta H)^2 + (0.154\Delta\delta N)^2\}^{1/2} \quad (1)$$

where  $\Delta\delta H$  and  $\Delta\delta N$  are the chemical shift differences in proton and nitrogen dimensions, respectively, and the weighting factor of 0.154 was used to scale the chemical shift differences of the  $^{15}\text{N}$  nuclei and the  $^1\text{H}$  nuclei (27). The chemical shift differences were then plotted onto the structure of LA5 (Figure 5D–F). This analysis reveals that

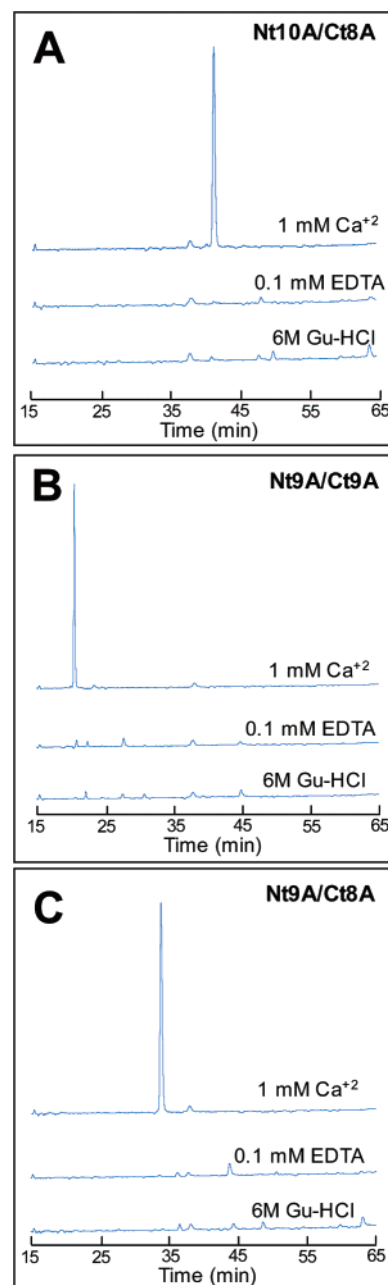


FIGURE 4: HPLC chromatograms of products formed after oxidative folding of Nt10A/Ct8A (A), Nt9A/Ct9A (B), and Nt9A/Ct8A (C). The proteins were incubated in a redox buffer containing calcium, EDTA, or denaturant as described in Experimental Procedures.

the alanine substitutions in the N-terminal lobe of Nt10A perturb residues in the C-terminal lobe involved either in calcium coordination or in van der Waals contact with the hairpin of the N-terminal lobe (Figure 5D). In contrast, the alanine substitutions in the C-terminal lobe of Ct9A affect the N-terminal lobe more broadly, with intermediate effects on chemical shift manifested at residues throughout the N-terminal lobe (Figure 5E). Finally, in the Nt9A/Ct9A mutant, which has alanine substitutions distributed across both lobes, the chemical shift perturbations observed are remarkably similar to those predicted by simple addition of the perturbations seen for the N-terminal and C-terminal mutants alone (Figure 5F).

*Calcium Affinity of Native and Mutant LA5 Modules.* Examination of the calcium dependence of folding (data not

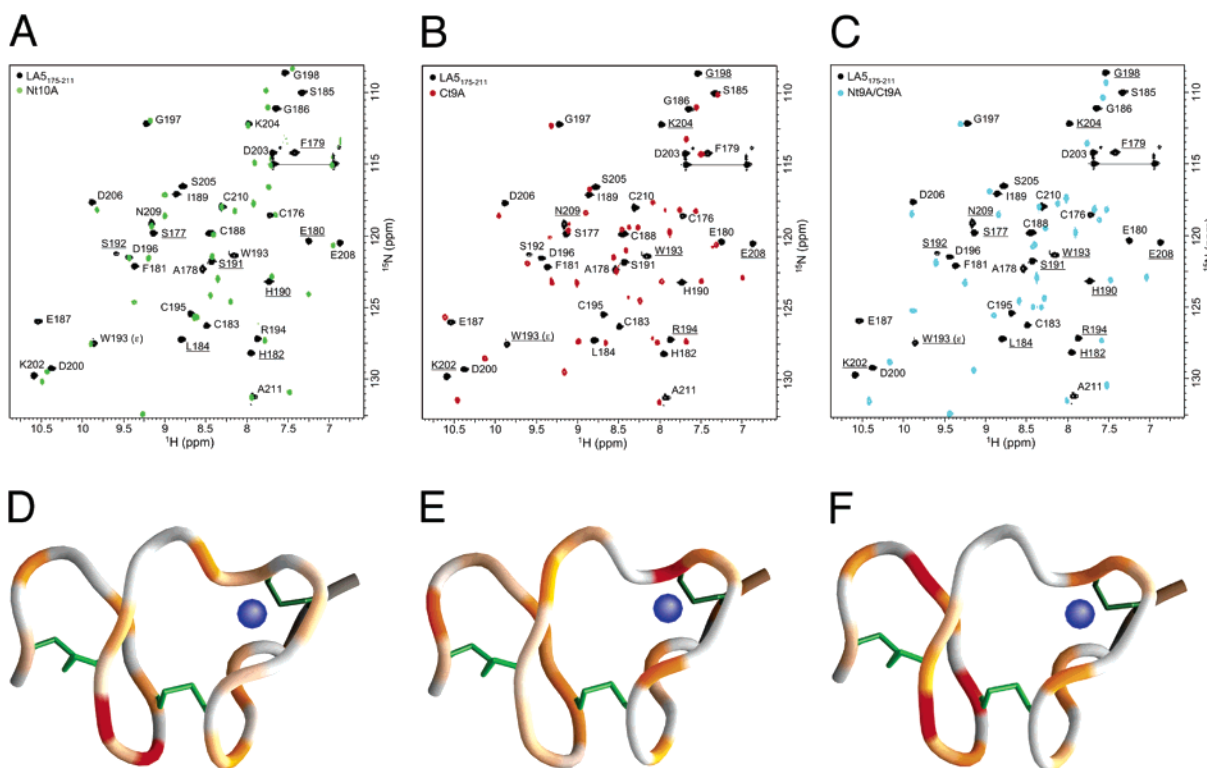


FIGURE 5: <sup>15</sup>N- and <sup>1</sup>H-HSQC NMR spectra (A–C) and chemical shift difference mapping (D–F) of Nt10A, Ct9A, and Nt9A/Ct9A in comparison to LA5<sub>[175–211]</sub>. (A) Superposition of the HSQC spectrum of Nt10A (green) onto that of native LA5<sub>[175–211]</sub> (black). (B) Superposition of the HSQC spectrum of Ct9A (red) onto that of native LA<sub>[175–211]</sub> (black). (C) Superposition of the HSQC spectrum of Nt9A/Ct9A (blue) onto that of native LA5<sub>[175–211]</sub> (black). In each panel, residue numbers denote backbone resonance assignments of native LA5<sub>[175–211]</sub>, and amino acid residues changed to Ala in the mutant are underlined. (D–F) Normalized chemical shift differences for Nt10A (D), Ct9A (E), and Nt9A/Ct9A (F) mapped onto a ribbon diagram of the LA5 structure. Normalized differences were calculated as described in Experimental Procedures, using the spectrum of LA5<sub>[175–211]</sub> as a reference. Individual residues of the ribbon are colored according to a sliding scale with gray (0), orange (0.5), and red (1.0) indicating negligible, intermediate, and maximal changes in chemical shift, respectively. Residues mutated to alanine were assigned a value of zero and are also colored gray. The disulfide bonds and calcium ion are depicted in green and blue, respectively.

shown) suggested that most of the mutants had a calcium affinity similar to that of LA5<sub>[175–211]</sub>. For LA5<sub>[175–211]</sub>, Nt10A, Nt9A, Ct8A, and Nt9A/Ct8A, we measured the calcium affinity directly by isothermal titration calorimetry at pH 5.6 (Figure 6). For LA5<sub>[175–211]</sub>, the measured affinity was  $990 \pm 30$  nM, an affinity consistent with other published values for LA5 obtained under a variety of different experimental conditions. The  $K_d$  values for the other mutants are within about an order of magnitude (Figure 6), showing that replacement of most of the nonconserved residues does not dramatically affect calcium binding. The Nt10A mutant has a measured  $K_d$  of  $2.4 \mu\text{M}$ , a little more than double that of LA5<sub>[175–211]</sub> at the experimental pH. Nt9A, Ct8A, and Nt9A/Ct8A exhibit a higher calcium affinity than LA5<sub>[175–211]</sub>, with the affinity of Nt9A/Ct8A equal to  $88 \pm 11$  nM.

**Folding of Native and Variant LA3–6 Minireceptors.** The existence of folding-competent LA5-derived domains with alanine-rich N-terminal and C-terminal lobes afforded the opportunity to dissect the importance of the nonconserved residues in each lobe with regard to apoE binding. Therefore, we made domain swaps in which LA5 was replaced by either Nt9A or Ct8A in the context of an LDLR minireceptor spanning LA3 through LA6 (LA3–6, see Figure 7A). These native and mutant minireceptors are denoted WT LA3–6, Nt9A–LA3–6, and Ct8A–LA3–6, respectively. Myc-tagged and untagged minireceptors all refold to a predomi-

nant isomer in the presence of calcium under conditions that permit disulfide exchange, as judged by reversed-phase HPLC (Figure 7B–G, top traces). In contrast, refolding in EDTA gives a broad distribution of isomers (Figure 7B–G, middle traces). By comparison, each minireceptor migrates as a single peak with a longer retention time after reduction with DTT (Figure 7B–G, bottom traces). The dominant disulfide-bonded isomer obtained upon folding in the presence of calcium was thus purified and used in apoE binding experiments (Figure 7B–G, top traces).

**Binding to DMPC-Complexed N–ApoE.** Binding of the different LA3–6 minireceptors to lipid-complexed apoE-(1–191) was compared in an immunoprecipitation assay using purified recombinant LA3–6 variants bearing a C-terminal myc epitope tag. Complexes were captured onto protein G beads loaded with an anti-myc antibody. Bound N–apoE–DMPC was then liberated by boiling in denaturing SDS loading buffer and detected by Western blot using an HRP-conjugated anti-apoE polyclonal antibody. The immunoprecipitation data (Figure 8A,B) confirm that the WT LA3–6 minireceptor binds N–apoE–DMPC in a specific, calcium-dependent manner, consistent with previous results (6). While Ct8A–LA3–6 also binds N–apoE–DMPC when calcium is present, Nt9A–LA3–6 does not (Figure 8A).

The specificity of these binding interactions was confirmed in a competition assay (Figure 8B). Both WT LA3–6 and



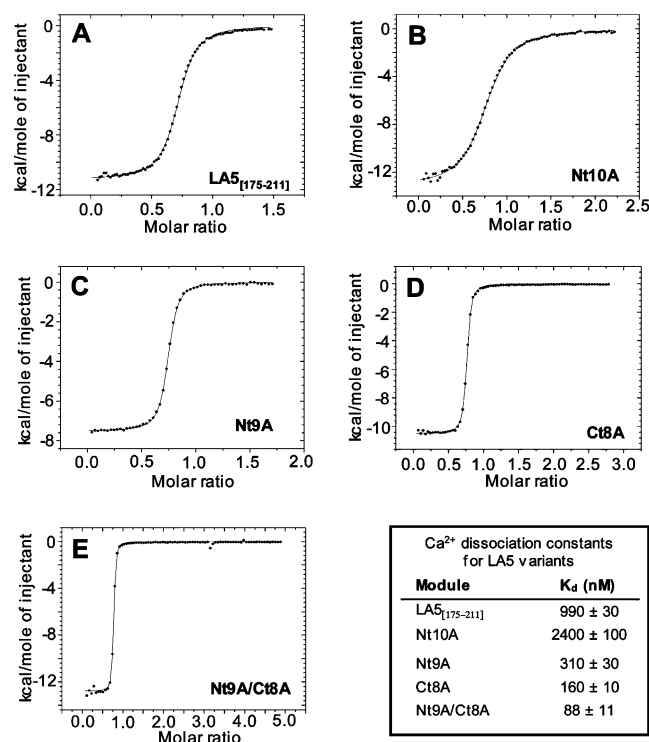


FIGURE 6: Calcium affinities of different LA modules determined by isothermal titration calorimetry. The calcium affinities of LA5<sub>[175-211]</sub> (A), Nt10A (B), Nt9A (C), Ct8A (D), and Nt9A/Ct8A (E) were measured in 20 mM MES buffer, pH 5.6. To calculate the calcium dissociation constant, data were fitted to a one-site binding model with the program Origin 5.0.

Ct8A–LA3–6 effectively compete with the binding of N–apoE–DMPC to WT LA3–6–myc, but Nt9A–LA3–6 does not, even when present in 20-fold molar excess. These results provide compelling evidence that Ct8A–LA3–6 maintains the capacity for high affinity binding to N–apoE–DMPC, while Nt9A–LA3–6 does not.

## DISCUSSION

Previous investigations using alanine saturation mutagenesis have probed sequence–structure relationships in a DNA-binding homeodomain, the Arc repressor, and the disulfide-bonded model protein BPTI (3–5), proteins in which secondary structure elements contribute significantly to directing the tertiary fold. In contrast, the relative paucity of secondary structure elements in LA modules offers a novel model for addressing folding determinants in disulfide-bonded structures that are not constrained by packing of their secondary structure elements.

Our “library” of alanine-rich LA5 variants identifies positions where the native side chains are not required to maintain the structural integrity of LA5 from the LDL receptor. Evidence that 7 of 8 mutants tested (Nt10A, Nt9A, Ct9A, Ct8A, Nt10A/Ct8A, Nt9A/Ct8A, and Nt9A/Ct9A) are natively folded comes from HPLC-based folding assays and solution NMR spectroscopy (Figures 2–5) showing calcium-dependent formation of a single well-folded isomer. ITC measurements of four mutants (Nt10A, Nt9A, Ct8A, and Nt9A/Ct8A) also demonstrate high-affinity binding of calcium in 1:1 stoichiometry, further confirming acquisition of the native fold (Figure 6). The proper folding of two mutants in which 18 of 37 residues (49%) are alanines echoes prior

findings from alanine saturation mutagenesis of BPTI, which retains a nativelike fold when at least 27 of its 58 residues (48%) are alanines (5).

The folding studies reported here define a set of “scaffold” residues sufficient to specify the LA fold. Among the key scaffolding residues are the conserved cysteines involved in disulfide bonding, the acidic residues participating in calcium coordination, two hydrophobic residues that pack against each other, two glycine residues that potentially confer flexibility, and other residues that form side chain intramolecular hydrogen bonds. Indeed, one of the glycines, Gly 197, lies in a position of the Ramachandran plot that is only favorable for glycine.

The increased calcium affinity of certain mutants (e.g., Nt9A/Ct8A) when compared with LA5<sub>[175-211]</sub> implies that the nonconserved residues of wild-type LA5 are not optimized for calcium binding affinity or domain stability. Replacement of two nonconserved lysine residues (Lys 202 and Lys 204) in the C-terminal lobe with alanines, for example, probably contributes to the increase in calcium affinity by rendering the electrostatic potential of the nearby calcium-coordinating site more negative.

Though most of the nonconserved residues can be replaced by alanine without disruption of the LA fold, consideration of sequence conservation alone does not completely predict which residues are required for folding, even for such a simple 40-residue module. Ser 185 and Glu 187, a pair of residues that form a hydrogen bond in the X-ray structure of LA5 (9), are not generally conserved among LA modules (Figure 1B,C). Nevertheless, replacement of Ser 185 and Glu 187 with alanines in the Nt12A mutant (in comparison to Nt10A) converts a folding-competent variant (Nt10A) into a folding-defective one (Nt12A). It may be the case that a larger sampling of LA modules sufficient to uncover covariation of the Ser 185–Glu 187 pair would have revealed the importance of these two residues in proper folding of the module.

Similarly, the significance of Glu 180 is not evident from analysis of consensus sequence alone. Nt10A differs from Nt9A solely by the additional replacement of Glu 180 with alanine, yet Nt10A binds calcium about 8-fold more weakly than Nt9A does (Figure 6). The reduction in calcium affinity resulting from the E180A substitution is qualitatively similar to that observed in an LA5 mutant lacking the N-terminal disulfide bond (22). The C176A/C188A LA5 variant has an unfolded N-terminal lobe but forms a nativelike C-terminal lobe that binds calcium with an affinity ~500-fold lower than wild-type LA5. As in the C176A/C188A case, the replacement of Glu 180 by alanine in Nt10A may destabilize or partially unlock the packing between the N- and C-terminal lobes, thereby indirectly reducing the calcium-binding affinity.

Comparison of the Ct9A and Ct10A mutants also reveals that Ser 205, which is highly conserved (89%) across the seed alignment, influences calcium affinity without directly participating in calcium coordination. Ct9A, which possesses Ser 205, has a calcium affinity approximately 6 times greater than that of wild-type LA5, yet Ct10A, which is identical to Ct9A except for the additional substitution of Ser 205 by alanine, only forms a preferred isomer when driven by the presence of calcium at higher concentration in the folding buffer (Figure 3C). Remarkably, the D203G mutation found



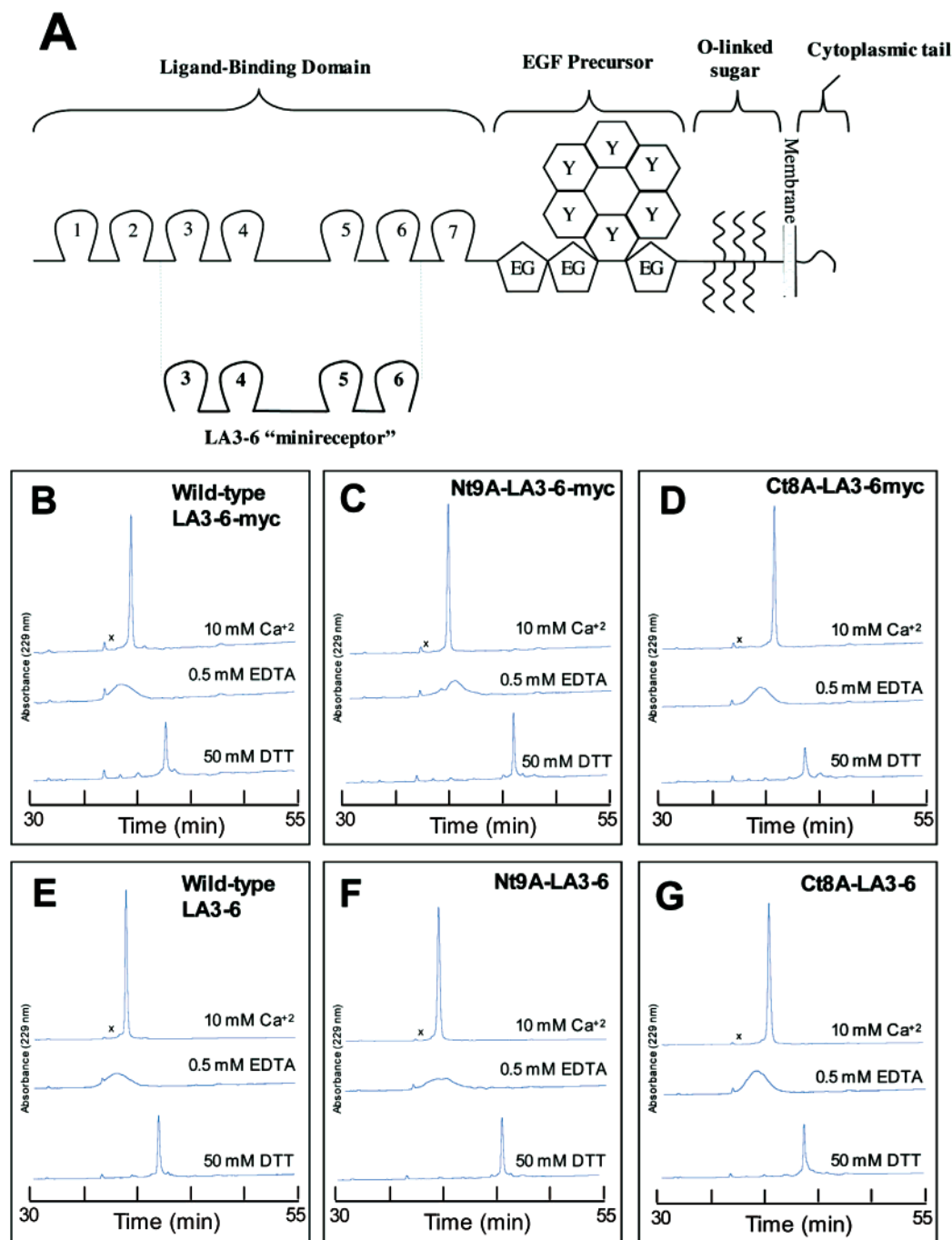


FIGURE 7: Purification of wild-type and mutant LA3-6 minireceptors for binding studies. (A) Modular organization of the LDLR. The truncated LA3-6 “minireceptor” used in this study is illustrated (EG, epidermal growth factor-like module; Y, repeat containing a YWTD motif). (B–G) HPLC chromatograms of purified minireceptors after (i) folding in a redox buffer with 10 mM  $\text{Ca}^{2+}$  (top traces), (ii) folding in a redox buffer with 0.5 mM EDTA (middle traces), or (iii) reduction with 50 mM DTT (bottom traces): (B) wild-type LA3-6-myc, (C) Nt9A-LA3-6-myc, (D) Ct8A-LA3-6-myc, (E) wild-type LA3-6, (F) Nt9A-LA3-6, (G) Ct8A-LA3-6. The “x” marked at an elution time of approximately 37 min denotes a column-derived impurity. The folded LA3-6 variants, which were analyzed in the chromatograms of the top traces, were used in the immunoprecipitation assays.

in FH patients, which alters the side chain hydrogen-bonding partner of Ser 205, also requires a roughly  $\sim 50$ -fold greater calcium concentration to fold to a preferred isomer (17), suggesting that this hydrogen-bonding network also stabilizes the calcium binding site.

In contrast to Glu 180 and Ser 205, Trp 193 does not appear to exhibit conservation (47%) primarily as a result of structural constraints. It may be that the conservation of tryptophan at this position reflects a functional rather than a structural constraint. Consistent with this possibility, the

interface between the LA4 and LA5 modules of the ligand-binding domain with the  $\beta$ -propeller, seen in the closed, low pH conformation of the LDLR, shows that both LA4 and LA5 have tryptophan residues at that position in the interface with the propeller (28).

**ApoE Binding Studies.** Previous studies have shown that LA5, the fifth ligand-binding repeat of the LDLR, is essential for binding of LDL,  $\beta$ -VLDL, and apoE-DMPC particles (29, 30). However, identification of specific residues at the contact interface has been elusive because almost all previous

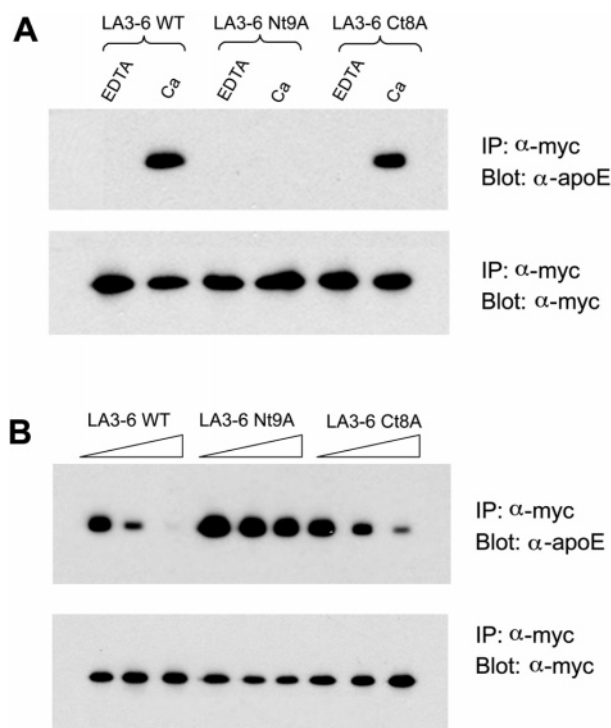


FIGURE 8: Binding of LA3-6 variants to N-apoE-DMPC in the presence and absence of calcium. Complexes were immunoprecipitated with anti-myc polyclonal antibody in the presence of either 2 mM EDTA or 2 mM CaCl<sub>2</sub> to recover complexes of wild-type and mutant myc-tagged minireceptors with N-apoE-DMPC. (A) Upper panel, Western blot was analyzed with an anti-apoE polyclonal antibody; bottom panel, after acidic stripping of the membrane, the same blot was analyzed by probing with an anti-myc monoclonal antibody. (B) Competition binding assay: immunoprecipitation with anti-myc antibody was carried out with the wild-type, myc-tagged LA3-6 in the presence of 1, 5, or 20 molar equiv of untagged competitor (wild-type LA3-6, Nt9A-LA3-6, or Ct8A-LA3-6). Upper panel, Western blot analyzed with an anti-apoE polyclonal antibody; bottom panel, after acidic stripping of the membrane, the same blot was probed with an anti-myc monoclonal antibody.

mutational studies altered conserved residues essential for module folding. As a consequence, any observed change in binding resulting from one of these mutations is likely to have resulted as an indirect consequence of module misfolding rather than from disruption of a site directly in contact with ligand (13, 17, 22, 31).

In these studies, we created chimeric LA3-6 minireceptors with multiple alanine substitutions in either the N- or C-terminal lobe of LA5 to investigate whether the replaced residues participate in binding of apoE-DMPC. The LA5 domain swaps were performed with Nt9A and Ct8A because these two variants fold cleanly to a unique disulfide isomer in the isolated module (Figures 2D, 3A), have a calcium affinity greater than native LA5 based on calorimetric measurements (Figure 6C,D), and permit folding of the LA3-6 chimeric minireceptors (Figure 7C,D,F,G). Though the immunoprecipitation assay is admittedly qualitative, the domain swap with Nt9A clearly disrupts apoE-DMPC binding, whereas the swap with Ct8A has little effect (Figure 8A,B). The domain swap experiment thus implicates nonconserved residues of the N-terminal lobe of LA5 in binding of apoE-DMPC and argues that nonconserved residues of the C-terminal lobe are unimportant in binding.

Among the residues replaced by alanine in the Nt9A mutant, His 190 and Ser 191 are most likely to participate directly in ligand binding. The importance of these two residues in binding was first suggested by the phenotype of a familial hypercholesterolemia mutation of the LDL receptor in which residues His 190 and Ser 191 are replaced by a single proline residue, because receptors with this mutation appear to be processed normally in transiently transfected cells but exhibit reduced LDL binding (32). These residues are also the only two residues of the N-terminal lobe in the intramolecular interface between LA5 and the  $\beta$ -propeller domain in the low-pH crystal structure (28), consistent with partial overlap of intermolecular (ligand-binding) and intramolecular interfaces and thus with the proposed intramolecular displacement model for low-pH-induced ligand release (28). The role of these residues, as well as other nonconserved residues in the N-terminal lobe, can now be directly addressed in future studies in which the side chains of native LA5 are reintroduced into the generic backbone of Nt9A. Among other residues seen in the intramolecular interface between LA5 and the  $\beta$ -propeller, Trp 193, Asp 196, and Asp 200 are conserved residues in the C-terminal lobe that are present in all domain-swapped minireceptors that were tested. These residues thus constitute an additional group of candidates that may potentially be part of an intermolecular, ligand-binding interface.

## ACKNOWLEDGMENT

We thank Sagar Koduri and Didem Vardar for helpful discussions and for critical review of the manuscript.

## REFERENCES

- Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonhammer, E. L., Studholme, D. J., Yeats, C., and Eddy, S. R. (2004) The Pfam protein families database, *Nucleic Acids Res.* 32, D138–41, (Database issue).
- Howell, B. W., and Herz, J. (2001) The LDL receptor gene family: signaling functions during development, *Curr. Opin. Neurobiol.* 11, 74–81.
- Shang, Z., Isaac, V. E., Li, H., Patel, L., Catron, K. M., Curran, T., Montelione, G. T., and Abate, C. (1994) Design of a "minimal" homeodomain: the N-terminal arm modulates DNA binding affinity and stabilizes homeodomain structure, *Proc. Natl. Acad. Sci. U.S.A.* 91, 8373–7.
- Brown, B. M., and Sauer, R. T. (1999) Tolerance of Arc repressor to multiple-alanine substitutions, *Proc. Natl. Acad. Sci. U.S.A.* 96, 1983–8.
- Kuroda, Y., and Kim, P. S. (2000) Folding of bovine pancreatic trypsin inhibitor (BPTI) variants in which almost half the residues are alanine, *J. Mol. Biol.* 298, 493–501.
- Fisher, C., Abdul-Aziz, D., and Blacklow, S. C. (2004) A two-module region of the low-density lipoprotein receptor sufficient for formation of complexes with apolipoprotein E ligands, *Biochemistry* 43, 1037–44.
- Daly, N. L., Djordjevic, J. T., Kroon, P. A., and Smith, R. (1995) Three-dimensional structure of the second cysteine-rich repeat from the human low-density lipoprotein receptor, *Biochemistry* 34, 14474–81.
- Daly, N. L., Scanlon, M. J., Djordjevic, J. T., Kroon, P. A., and Smith, R. (1995) Three-dimensional structure of a cysteine-rich repeat from the low-density lipoprotein receptor, *Proc. Natl. Acad. Sci. U.S.A.* 92, 6334–8.
- Fass, D., Blacklow, S., Kim, P. S., and Berger, J. M. (1997) Molecular basis of familial hypercholesterolemia from structure of LDL receptor module, *Nature* 388, 691–3.

10. Huang, W., Dolmer, K., and Gettins, P. G. (1999) NMR solution structure of complement-like repeat CR8 from the low density lipoprotein receptor-related protein, *J. Biol. Chem.* 274, 14130–6.
11. Clayton, D., Brereton, I. M., Kroon, P. A., and Smith, R. (2000) Three-dimensional NMR structure of the sixth ligand-binding module of the human LDL receptor: comparison of two adjacent modules with different ligand binding specificities, *FEBS Lett.* 479, 118–22.
12. Dolmer, K., Huang, W., and Gettins, P. G. (2000) NMR solution structure of complement-like repeat CR3 from the low density lipoprotein receptor-related protein. Evidence for specific binding to the receptor binding domain of human alpha(2)-macroglobulin, *J. Biol. Chem.* 275, 3264–9.
13. North, C. L., and Blacklow, S. C. (2000) Solution structure of the sixth LDL-A module of the LDL receptor, *Biochemistry* 39, 2564–71.
14. Tonelli, M., Peters, R. J., James, T. L., and Agard, D. A. (2001) The solution structure of the viral binding domain of Tva, the cellular receptor for subgroup A avian leukosis and sarcoma virus, *FEBS Lett.* 509, 161–8.
15. Simonovic, M., Dolmer, K., Huang, W., Strickland, D. K., Volz, K., and Gettins, P. G. (2001) Calcium coordination and pH dependence of the calcium affinity of ligand-binding repeat CR7 from the LRP. Comparison with related domains from the LRP and the LDL receptor, *Biochemistry* 40, 15127–34.
16. Wang, Q. Y., Huang, W., Dolmer, K., Gettins, P. G., and Rong, L. (2002) Solution structure of the viral receptor domain of Tva and its implications in viral entry, *J. Virol.* 76, 2848–56.
17. Blacklow, S. C., and Kim, P. S. (1996) Protein folding and calcium binding defects arising from familial hypercholesterolemia mutations of the LDL receptor, *Nat. Struct. Biol.* 3, 758–62.
18. Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins, *Biochemistry* 6, 1948–54.
19. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Sci.* 4, 2411–23.
20. Pons, J.-L., Malliavin, T. E., and Delsuc, M. A. (1996) Gifa V.4: a complete package for NMR data set processing, *J. Biomol. NMR* 8, 445–52.
21. Cavanagh, J., Palmer, I. A. G., Fairbrother, W., and Skelton, N. J. (1996) *Protein NMR Spectroscopy. Principles and Practice*, Academic Press: San Diego, CA.
22. Koduri, V., and Blacklow, S. C. (2001) Folding determinants of LDL receptor type A module, *Biochemistry* 40, 12801–7.
23. Schuster-Bockler, B., Schultz, J., and Rahmann, S. (2004) HMM Logos for visualization of protein families, *BMC Bioinf.* 5, 7.
24. Atkins, A. R., Brereton, I. M., Kroon, P. A., Lee, H. T., and Smith, R. (1998) Calcium is essential for the structural integrity of the cysteine-rich, ligand-binding repeat of the low-density lipoprotein receptor, *Biochemistry* 37, 1662–70.
25. Dolmer, K., Huang, W., and Gettins, P. G. (1998) Characterization of the calcium site in two complement-like domains from the low-density lipoprotein receptor-related protein (LRP) and comparison with a repeat from the low-density lipoprotein receptor, *Biochemistry* 37, 17016–23.
26. Wang, Q. Y., Dolmer, K., Huang, W., Gettins, P. G., and Rong, L. (2001) Role of calcium in protein folding and function of Tva, the receptor of subgroup A avian sarcoma and leukosis virus, *J. Virol.* 75, 2051–8.
27. Seavey, B. R., Farr, E. A., Westler, W. M., and Markley, J. L. (1991) A relational database for sequence-specific protein NMR data, *J. Biomol. NMR* 1, 217–36.
28. Rudenko, G., Henry, L., Henderson, K., Ichtchenko, K., Brown, M. S., Goldstein, J. L., and Deisenhofer, J. (2002) Structure of the LDL receptor extracellular domain at endosomal pH, *Science* 298, 2353–8.
29. Esser, V., Limbird, L. E., Brown, M. S., Goldstein, J. L., and Russell, D. W. (1988) Mutational analysis of the ligand binding domain of the low density lipoprotein receptor, *J. Biol. Chem.* 263, 13282–90.
30. Russell, D. W., Brown, M. S., and Goldstein, J. L. (1989) Different combinations of cysteine-rich repeats mediate binding of low density lipoprotein receptor to two different proteins, *J. Biol. Chem.* 264, 21682–8.
31. Djordjevic, J. T., Bieri, S., Smith, R., and Kroon, P. A. (1996) A deletion in the first cysteine-rich repeat of the low-density-lipoprotein receptor leads to the formation of multiple misfolded isomers, *Eur. J. Biochem.* 239, 214–9.
32. Sun, X. M., Patel, D. D., Webb, J. C., Knight, B. L., Fan, L. M., Cai, H. J., and Soutar, A. K. (1994) Familial hypercholesterolemia in China. Identification of mutations in the LDL-receptor gene that result in a receptor-negative phenotype, *Arterioscler. Thromb.* 14, 85–94.

BI047575J