Solution Structure of the Sixth LDL-A Module of the LDL Receptor[†,‡]

Christopher L. North and Stephen C. Blacklow*

Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, 75 Francis Street, Boston, Massachusetts 02115

Received September 7, 1999; Revised Manuscript Received November 29, 1999

ABSTRACT: The low-density lipoprotein receptor (LDLR) is the primary mechanism for uptake of plasma cholesterol into cells and serves as a prototype for an entire class of cell surface receptors. The amino-terminal domain of the receptor consists of seven LDL-A modules; the third through the seventh modules all contribute to the binding of low-density lipoproteins (LDLs). Here, we present the NMR solution structure of the sixth LDL-A module (LR6*) from the ligand binding domain of the LDLR. This module, which has little recognizable secondary structure, retains the essential structural features observed in the crystal structure of LDL-A module five (LR5) of the LDLR. Three disulfide bonds, a pair of buried residues forming a hydrophobic "mini-core", and a calcium-binding site that serves to organize the C-terminal lobe of the module all occupy positions in LR6* similar to those observed in LR5. The striking presence of a conserved patch of negative surface electrostatic potential among LDL-A modules of known structure suggests that ligand recognition by these repeats is likely to be mediated in part by electrostatic complementarity of receptor and ligand. Two variants of LR6*, identified originally as familial hypercholesterolemia (FH) mutations, have been investigated for their ability to form native disulfide bonds under conditions that permit disulfide exchange. The first, E219K, lies near the amino-terminal end of LR6*, whereas the second, D245E, alters one of the aspartate side chains that directly coordinate the bound calcium ion. After equilibration at physiologic calcium concentrations, neither E219K nor D245E folds to a unique disulfide isomer, indicating that FH mutations both within and distant from the calcium-binding site give rise to protein-folding defects.

The low-density lipoprotein receptor (LDLR)[1] is the primary mechanism for uptake of plasma cholesterol into cells (*1*) and serves as a prototype for an entire class of cell surface receptors (see ref *2* for a recent review). This family of cell surface receptors includes the low-density lipoprotein receptor-related protein (*3*), megalin/gp330 (*4*), and the very low-density lipoprotein receptors (*5*, *6*). All of these receptors share a similar modular organization (Figure 1) and all use multiple LDL-A modules to bind their ligands.

The amino-terminal domain of the LDLR consists of seven LDL-A modules (*7*), which are responsible for the binding of lipoproteins. Each LDL-A module spans about 40 amino acids and contains six cysteine residues engaged in three disulfide bonds (Figure 1). Modules 3−7 contribute to the binding of low-density lipoproteins (LDLs), whereas the greatest contribution to the binding of β-migrating very low-density lipoproteins (β-VLDLs) comes from module 5 (*8*).

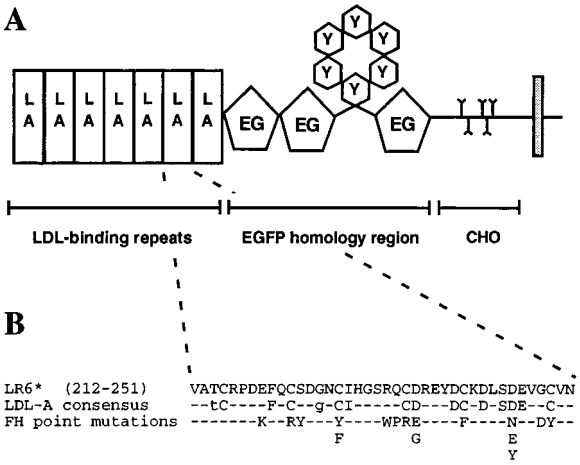

FIGURE 1: (A) Schematic diagram of the modular arrangement of the LDL receptor: LA, LDL-A ligand-binding modules; EG, EGF-like modules; Y, YWTD repeats; CHO, O-glycosylation region. Adapted from ref *41*. (B) Amino acid sequence of LR6*. Consensus residues are in uppercase when conserved in at least six of the seven LDL-A modules of the LDLR, and in lowercase when conserved in five of the seven modules. Point mutations that result in familial hypercholesterolemia, listed in the LDLR Locus database at http://www.ucl.ac.uk/fh/ (*42*), are indicated at the bottom.

Familial hypercholesterolemia (FH), characterized clinically by marked elevation in plasma LDL and cholesterol

[†] Supported by NIH Grant HL61001-01 (to S.C.B.). This work was conducted while S.C.B. was a Pfizer Scholar. S.C.B. is a Pew Scholar in the Biomedical Sciences.

[‡] The coordinates of LR6* are available from the Protein Data Bank under file name 1d2j.

* To whom correspondence should be addressed. Telephone: (617) 732-5799. Fax: (617) 264-5296. E-mail: sblacklow@rics.bwh.harvard.edu.

[1] Abbreviations: 2D, two-dimensional; 3D, three-dimensional; ApoE, apolipoprotein E; β-VLDL, β-migrating very low-desnity lipoprotein; FH, familial hypercholesterolemia; GSH, reduced glutathione; GSSG, oxidized glutathione; HSQC, heteronuclear single-quantum coherence; LDL, low-density lipoprotein; LDLR, low-density lipoprotein receptor; LR5, LDL-A module 5 of the LDL receptor; LR6, LDL-A module 6 of the LDL receptor; LR6*, LDL-A module 6 of the LDL receptor, with the M243L substitution; NOE, nuclear Overhaüser effect; TOCSY, total correlation spectroscopy.

10.1021/bi992087a CCC: $19.00
Published on Web 02/16/2000

levels, results from mutations in the LDLR (*9*). Heterozygotes, numbering about one in 500 persons, have a substantially increased risk for coronary artery disease, and homozygous FH leads to death from coronary atherosclerosis at a very young age. Many of the FH alleles are point mutations scattered among the LDL-A repeats (http://www.ucl.ac.uk/fh/).

The 1.7 Å crystal structure (*10*) of LDL-A module 5 of the LDL receptor (LR5) provides an explanation for the observed folding defects (*11*) that result from several point mutations within LR5 known to cause FH. The structure of LR5 contains a calcium ion coordinated in octahedral geometry by acidic residues that lie at the carboxy terminus of the module and that are conserved among LDL-A repeats. The three disulfide bonds constrain the overall module topology, in which the only elements of recognizable secondary structure include two short antiparallel β-strands and a single turn of 3_{10}-helix. FH mutations in LR5 alter the side chains of residues that either directly coordinate calcium or serve as scaffolding residues of the module (*10*).

Here we present the solution structure of module 6 (LR6*) of the LDL receptor determined by NMR spectroscopy. This module retains the essential structural features observed in the crystal structure of LDL-A module 5 of the LDLR. The presence of a conserved patch of negative surface electrostatic potential among LDL-A modules of known structure suggests that ligand recognition by these repeats is likely to be mediated in part by electrostatic complementarity of receptor and ligand.

Under conditions permitting disulfide exchange, two FH-causing mutations, one (E219K) near the N-terminus of LR6 and the other (D245E) in the C-terminal calcium-binding site, fail to fold efficiently to the native disulfide isomer. These folding defects are rationalized on the basis of the LR6* solution structure, which also explains why other mutations in LR6 might interfere with proper module folding.

MATERIALS AND METHODS

Protein Expression and Purification. The designation LR6* is used throughout to indicate the presence of an M243L substitution, which has been introduced into the native LR6 sequence to simplify protein expression and purification. Unlabeled and ^{15}N-labeled LR6* were expressed in the BL21(DE3) plys(S) cell line using the plasmid pMM-LR6* and were purified as described previously (*12*). The E219K and D245E mutations of LR6* were constructed by oligonucleotide-directed mutagenesis in the CJ236 cell line (*13*), and the identity of each mutant was confirmed by DNA sequencing. Expression and purification of these mutants was achieved using the procedure employed to produce LR6* (*12*).

NMR Spectroscopy. All NMR spectra were acquired on either a Varian Unity spectrometer at 11.75 T or a Varian Unity Plus spectrometer at 9.4 T. Spectra were processed and analyzed using the programs Felix 97 (MSI) and XEASY (*14*). Lyophilized samples of LR6* were dissolved into a solution of either 90% H_2O/10% D_2O or 99.8% D_2O containing 10 mM $CaCl_2$. The measured pH of the sample was adjusted to a value of 5.2 with NaOH; the final protein concentration was ~2 mM, and all experiments were performed at 298 K, except as noted. Resonance assignments were determined primarily from ^{15}N-HSQC (*15*), ^{15}N-NOESY-HSQC [512 × 128 × 32 complex points, 150 ms mixing time, WATERGATE (*16*) water suppression, 11.75 T; *17*], and ^{15}N-TOCSY-HSQC (512 × 128 × 32 complex points, 60 ms mixing time, 11.75 T; *17*). Additional frequency assignments were determined using 2D-TOCSY (1024 × 256 complex points, 60 ms mixing time, WATERGATE water suppression, 11.75 T), 2D-NOESY (150 ms mixing time, WATERGATE water suppression, 11.75 T), and 2D-P-COSY (9.4 T; *18*) spectra.

Measurement of Hydrogen Exchange Rates. To measure slow hydrogen exchange rates, a ^{15}N-labeled protein sample was dissolved in an aqueous (H_2O) solution containing 10 mM $CaCl_2$. After the pH was adjusted to 5.2 with NaOH, the sample was lyophilized. Hydrogen exchange was initiated at 25 °C by dissolving this lyophilized sample in a solution of 99% D_2O. Exchange rates for individual amide protons were calculated from ^{15}N-HSQC spectra (512 × 32 complex points with four scans per block) by fitting the decay of integrated peak volumes with time. To determine the protection factors reported in Figure 2D, the intrinsic hydrogen exchange rates were predicted as described by Englander and colleagues (*19*) using the program Sphere (http://dino.fold.fccc.edu:8080).

Structure Calculations. NOESY peaks were picked, assigned, and integrated using the program XEASY, version 1.3.13. Assignments of NOESY spectra were refined in DYANA using the ASNO program (*20*) followed by manual review in XEASY. NOE distance constraints were primarily derived from ^{15}N-NOESY-HSQC (512 × 256 × 32 complex points, mixing times of 50 and 150 ms, WATERGATE water suppression, 11.75 T) and 2D-NOESY (1024 × 512 complex points, mixing times of 50 and 150 ms, WATERGATE water suppression, 11.75 T) spectra. Additional NOE constraints were derived from a 2D-NOESY spectrum in 100% D_2O (1024 × 512 complex points, 150 ms mixing time, 9.4 T). J_{HNHA} couplings for dihedral angle constraints were measured from a ^{15}N-HMQC-J (1024 × 512 complex points, 9.4 T) spectrum (*21*). Distance constraints derived from the 50 ms NOESY spectra were calibrated in DYANA (*22*) using the program CALIBA (*23*). NOE resonances from the 150 ms mixing time NOESY spectra were classified as strong, medium, weak, and very weak and were given generous upper limits of 3.3, 3.8, 4.5, and 5.5 Å, respectively. Pseudoatom corrections were included when stereospecific assignments were not available.

Preliminary structures were calculated using DYANA with disulfide bond geometry defined by upper and lower distance limits. Redundant constraints were kept throughout the DYANA calculations to add weight to NOE peaks which appeared in more than one spectrum. After it became apparent that the backbone trace of LR6* in the region of the calcium-binding site was highly similar to that of LR5, distance constraints were created to define an octahedral geometry for the calcium-binding site of LR6*. From sequence alignment, the oxygen atoms analogous to those coordinating calcium in the LR5 structure were constrained to specify an octahedral geometry around the bound calcium atom in LR6*. Specifically, the octahedral geometry of the calcium-binding site was defined by placing an upper distance limit of 5.0 Å between each of the three opposite pairs of oxygen atoms and by placing bounds of 3.3−3.7 Å

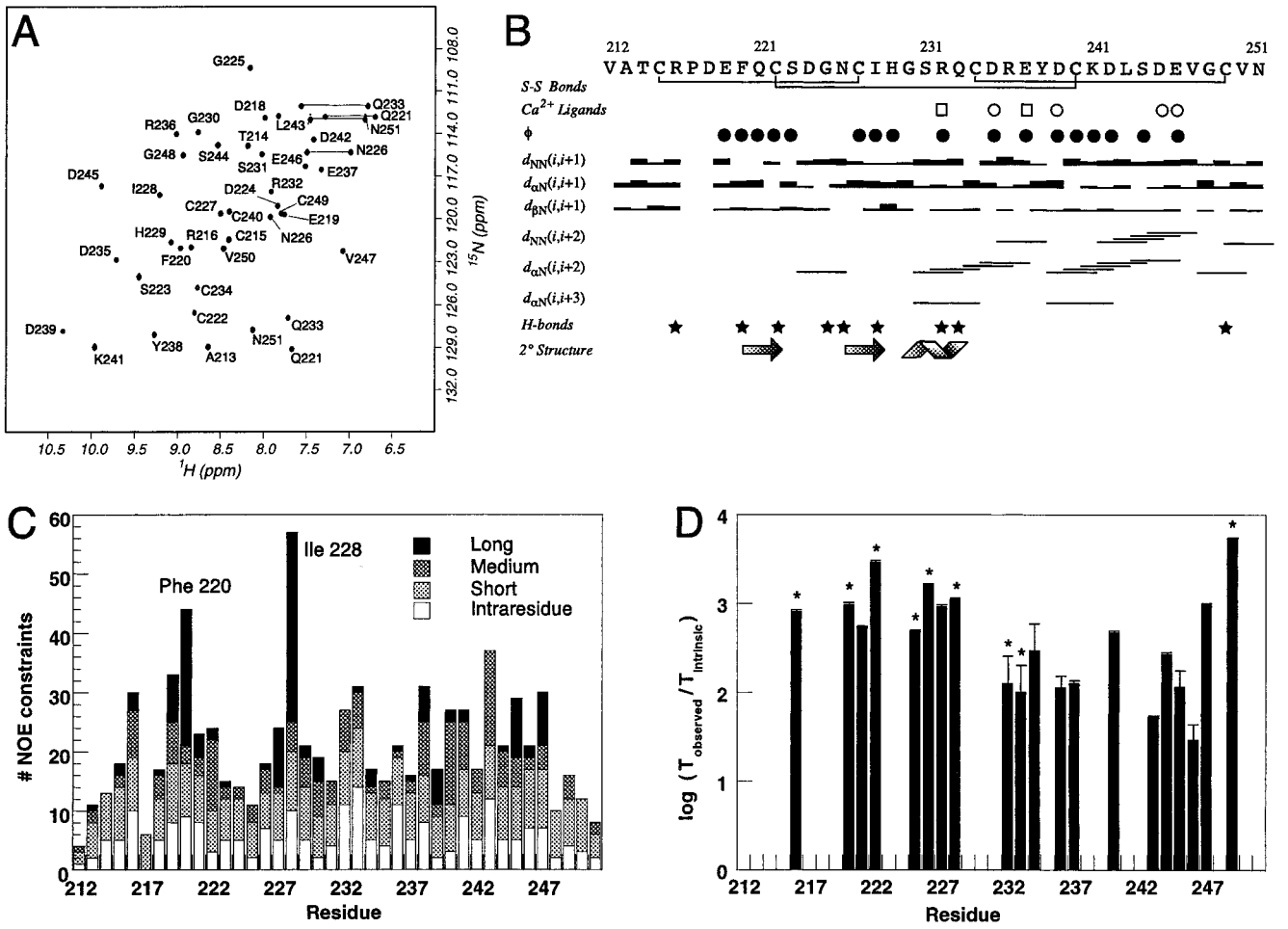

FIGURE 2: (A) ^{15}N−^{1}H HSQC spectrum of LR6* with assignments. (B) Summary of local constraints used in deriving the structure of LR6*, including disulfide bonds, calcium coordinating residues (□ indicates backbone carbonyl oxygen, ○ indicates side chain carboxylate oxygen), hydrogen bonding backbone amides, J_{HNHA}-derived constraints, and short- to medium-range NOE-derived backbone constraints. (C) Plot of NOE constraints per residue. As the plot indicates, Phe 220 and Ile 228 have the greatest number of long-range constraints. (D) Protection factors of amide protons from LR6*, estimated from hydrogen exchange data. Protected amides for which hydrogen-bonding partners were specified in structure calculations are denoted with an asterisk.

between neighboring pairs of oxygen atoms. Hydrogen bonding constraints consistent with preliminary calculated structures and hydrogen exchange data were added in the final stages of refinement (Table 1).

The best 30 of 60 model structures from DYANA (i.e., those with the lowest target function values) were chosen for final refinement in XPLOR (*24*). Before refinement, an explicit calcium ion was placed into each model at the midpoint between coordinating backbone carbonyl oxygen atoms. The 5.0 Å upper distance limits between the three opposite pairs of coordinating oxygen atoms were then replaced by six 2.5 Å upper distance limits between each of the coordinating oxygen atoms and the calcium ion. Two hundred steps of energy minimization were then performed to relieve resulting van der Waals conflicts. Disulfide bonds were considered as covalent bonds in XPLOR. Refinement in XPLOR followed the refine.inp protocol (*24*). The 20 models with the lowest total energy values are those reported here. The energy-minimized mean structure (Figures 3−5) was also calculated with XPLOR. PROCHECK-NMR and AQUA (*25*) were used to assess structure quality and fit to NOE restraints [the Supporting Information contains ribbon diagrams of structures calculated when H-bonds, calcium restraints, or both H-bonds and calcium restraints are omitted

Table 1: Constraint and Structural Statistics

Constraint Statistics	
no. of NOE distances	
intraresidue ($r = 0$)	232
sequential ($r = 1$)	155
medium-range ($1 < r < 4$)	85
long-range ($r > 4$)	69
total	541
additional constraints	
no. of disulfide bonds	3
no. of H-bonds	9
no. of calcium site distance constraints	18
J_{HNHA}-derived ϕ angles	17
Structural Statistics	
coordinate precision (residues 215−249)	
rmsd to mean (backbone atoms)	0.55 Å
rmsd to mean (heavy atoms)	1.01 Å
PROCHECK statistics (residues 215−249)	
no. of residues in favored regions	50%
no. of residues in additional allowed regions	35%
restraint satisfaction	
average no. of NOE violations	56
maximum NOE violation	0.37 Å
rms violation	0.07 ± 0.06 Å
XPLOR energies (kcal/mol)	
average E_{total}	162
average E_{NOE}	40

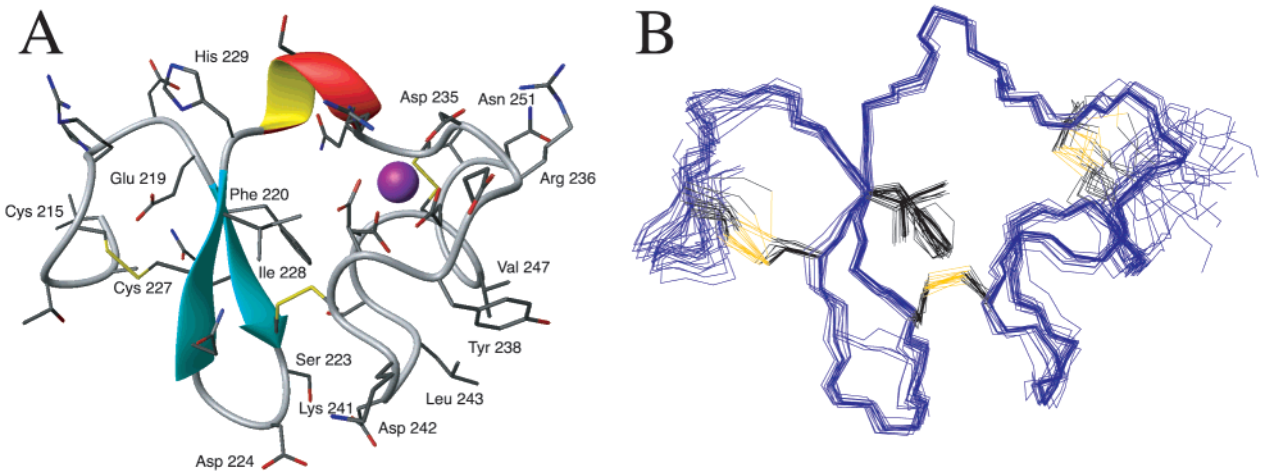

FIGURE 3: (A) Energy-minimized mean structure of LR6*, with selected residues labeled as reference points. The position of the calcium ion is indicated by a purple sphere. (B) Superposition of the 20 lowest-energy final calculated structures. Backbone atoms (blue) are illustrated along with side chains of the isoleucine, phenylalanine, and cysteine residues (black). Disulfide bonds are in yellow. This figure was prepared with the program MOLMOL (*43*, *44*).

from the structure calculations (Figure 1 of the Supporting Information)]. Electrostatic maps were calculated with the program GRASP (*26*), using the CHARMM partial charge and radius parameters.

Disulfide Exchange Experiments. The calcium dependence of folding was investigated for each peptide under conditions permitting disulfide exchange. LR6*, E219K, or D245E (10 μM) was allowed to rearrange in an anaerobic chamber (Coy Laboratory Products) at room temperature in a redox buffer of 10 mM Tris (pH 8.5) containing 2.5 mM reduced glutathione (GSH), 0.5 mM oxidized glutathione (GSSG), and either 1 mM $CaCl_2$ or 1 mM EDTA.

Aliquots were removed from each sample at 24 and 48 h; disulfide exchange was stopped by addition of acetic acid to a final concentration of 5% (v/v). Samples were analyzed by reversed-phase HPLC on a Vydac C-18 column using a linear gradient of 0.1% of solvent B/min. Solvent reservoirs contained water with 0.1% trifluoroacetic acid (A) and 90% acetonitrile with 0.1% trifluoroacetic acid (B). Comparison of 24 and 48 h time points suggests that an equilibrium distribution of disulfide-bonded isomers is reached within 24 h.

RESULTS

Assignments, Constraints, and Structure Calculations. The HSQC spectrum of LR6* exhibits good chemical shift dispersion in the presence of calcium, giving a single set of resonances consistent with a unique, well-defined conformation (Figure 2A). All assignments have therefore been derived from spectra acquired in the presence of saturating concentrations of calcium ion.

Constraint and structural statistics are summarized in Table 1. Sequential and medium-range NOE restraints along with exchange-protected backbone amides are illustrated in Figure 2B. The greatest number of long-range NOE constraints arises from residues Phe 220 and Ile 228 which form a small central core of the module (Figure 2C). A total of 17 dihedral angle constraints were derived from measurements of $^3J_{HNHA}$ coupling constants.

A total of 18 backbone amide protons are protected against hydrogen exchange in LR6* (Figure 2D). Hydrogen bonding partners were seen for nine amides in the initial structure calculations, and explicit constraints for these amides were added in the final refinement steps. Backbone amides of residues 220−222 and 225−228, which define a short antiparallel β-sheet, are protected against exchange, as are the backbone amides of residues 232 and 233, which are the N-terminal residues of the 3_{10}-helix. Backbone amides of residues 216 and 249 form hydrogen bonds across loops near the N- and C-termini of LR6. Residues 243−247 in the tightly packed calcium-binding region are also protected.

Description of the LR6 Structure*. The 20 lowest-energy structures of LR6* represent a single conformation (Table 1), with well-defined backbone coordinates between the first and last cysteine residues of the module (Figure 3). After final refinement in XPLOR, none of the models contain violations of the experimental distance constraints of greater than 0.37 Å (Table 1). None of the dihedral angle constraints are violated by more than 5° in any of the final models. Of the well-ordered residues (215−249), only Lys 241 falls in a disallowed region of the Ramachandran plot.

Because the calcium-binding site of LR6* is highly specific for calcium and does not accommodate other divalent metals (S. C. Blacklow, unpublished results; see also ref *27*), the geometry of the calcium-binding site of LR6* has been modeled on the basis of the coordination site observed in the LR5 crystal structure. From sequence alignment, the oxygen atoms analogous to those coordinating calcium in the LR5 structure have been constrained to specify an octahedral geometry around the bound calcium atom in LR6*. Thus, the carboxylate oxygens from D235, D239, D245, and E246 occupy four of the coordination sites, and the final two coordination sites are occupied by the backbone carbonyl oxygens of residues R232 and E237 (Figure 4B; see Materials and Methods).

We considered the possibility that the carbonyl group of Tyr 238 (rather than that of Glu 237) participates directly in coordination of the calcium ion. When a simulated annealing model was calculated with the tyrosine carbonyl oxygen as a calcium ligand, an acceptable calcium-binding site was formed, all NOE constraints were accommodated within 0.5 Å, and the tyrosine ring was placed in position to account

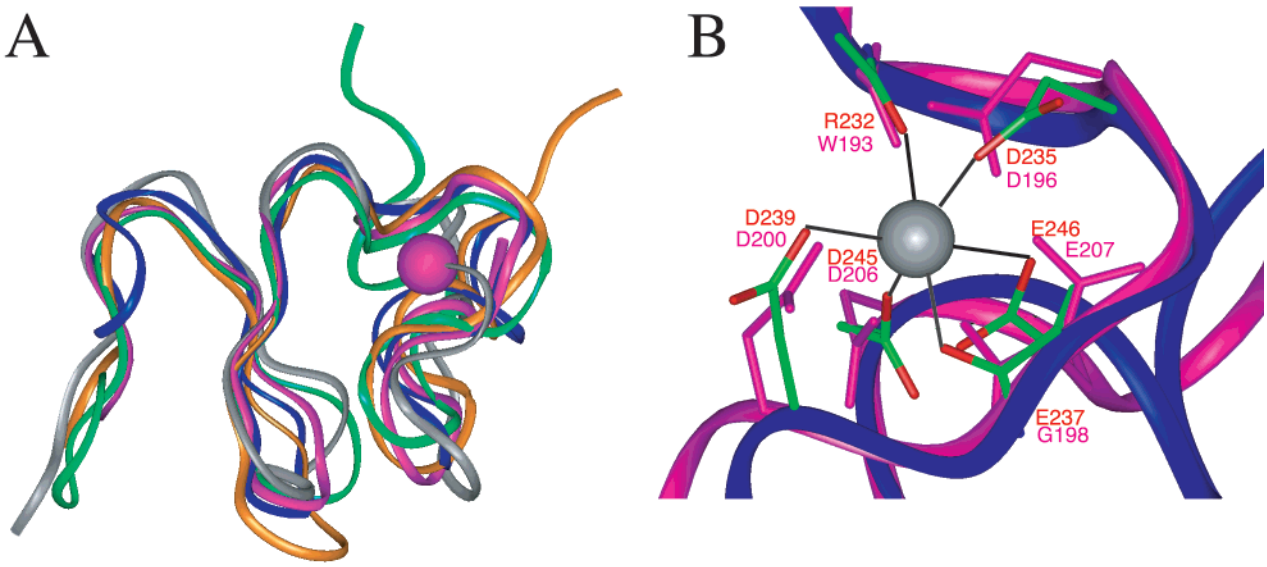

FIGURE 4: (A) Comparison of backbone atomic coordinates for all LDL-A modules with published structures: green, LR1 of the LDLR (*33*); gray, LR2 of the LDLR (*34*); pink, LR5 of the LDLR (*10*); orange, CR8 of the low-density lipoprotein receptor-related protein (*28*); and blue, LR6*. (B) Comparison of the Ca^{2+} binding sites of LR6* (colored by atom type) and LR5 (pink). Coordinates of the structures used for comparison were obtained from the Protein Data Bank (LR1, 1ldl; LR2, 1ldr; LR5, 1ajj; CR8, 1cr8). Structural comparisons were performed using the program Insight II (MSI).

for the observed upfield shift of the Leu 243 resonances. However, this alternative model resulted in a strained local geometry, in which both the XPLOR total energy increased by >30 kcal and the ϕ and ψ angles of Tyr 238 were placed in a disallowed region of the Ramachandran map. Furthermore, the three-residue spacing between the conserved aspartates of the sequence CDxZxDC (where Z corresponds to the residue from which the calcium-coordinating carbonyl group is derived in LR5 and CR8; *28*) is itself highly conserved, suggesting that the calcium ligand is always the carbonyl group from the central amino acid between the aspartates. For these reasons, we strongly favor the model in which the carbonyl group of Glu 237 is the calcium-coordinating ligand.

Folding of Modules with FH Mutations. In the presence of calcium under conditions permitting disulfide exchange, wild-type LR6* folds predominantly to the native disulfide isomer, whereas a distribution of non-native isomers predominate in the absence of calcium under otherwise identical conditions (*12*).

Two FH mutations, E219K and D245E, were introduced into LR6*, and were examined for the ability to fold to the native disulfide-bonded isomer. Under conditions that result in near-quantitative folding of LR6*, neither E219K nor D245E folds to a predominant disulfide isomer after 48 h (Figure 6). Indeed, the distribution of isomers is similar, regardless of whether folding is attempted in the presence of calcium (1 mM), in the absence of calcium (1 mM), or in denaturant (6 M guanidine·HCl; data not shown). Similar results have been found previously for FH mutations in LDL-A modules 5 (*11*) and 1 (*29*).

DISCUSSION

In total, more than 100 proteins that contain LDL-A modules are present in the sequence database, including not only members of the LDL receptor family but also components of the complement cascade (*30*), G protein-coupled receptors (*31*), and the receptor for Rous sarcoma virus type A (*32*). Principles that underlie ligand recognition by the LDL-A modules of the LDLR are likely to apply to the broad range of family members that contain a similar composition of extracellular modules.

Overview of LR6 Structure*. The backbone of LR6* is well-defined for all residues between the first and last cysteines of the module (Figure 3), and LR6* retains the essential structural features observed in the crystal structure of LR5 (*10*) and in the NMR structure of CR8 (*28*). These conserved structural features include three disulfide bonds, a short antiparallel β-sheet, a single turn of 3_{10}-helix, and a high-affinity calcium-binding site. The same global topology is also seen in the NMR structures of LR1 (*33*) and LR2 (*34*), the structures of which were determined in solution without explicit consideration of the calcium-binding site (Figure 4A).

The LR6* structure (Figure 3) consists of two loops, which define N- and C-terminal "lobes" of the molecule, connected by the 3_{10}-helix at one end and by a disulfide bond between the second and fifth conserved cysteine residues at the other. The antiparallel β-sheet, together with a disulfide bond between the first and third conserved cysteines, serves as a scaffold for the N-terminal lobe of the module. The C-terminal lobe is organized around the bound calcium ion (Figure 4B) and is also held together by a conserved disulfide bond between the fourth and sixth cysteines. The coordination sphere of this bound calcium ion is well-modeled in octahedral geometry, with four carboxylate oxygen atoms from the conserved acidic motif in one plane and the two carbonyl oxygen atoms completing the coordination sphere at the apices.

The short antiparallel β-sheet positions the conserved hydrophobic residues (Phe 220 and Ile 228) into the interior of the module. These conserved residues occupy similar positions in the crystal structure of LR5, suggesting a critical scaffolding role for these side chains in stabilizing the

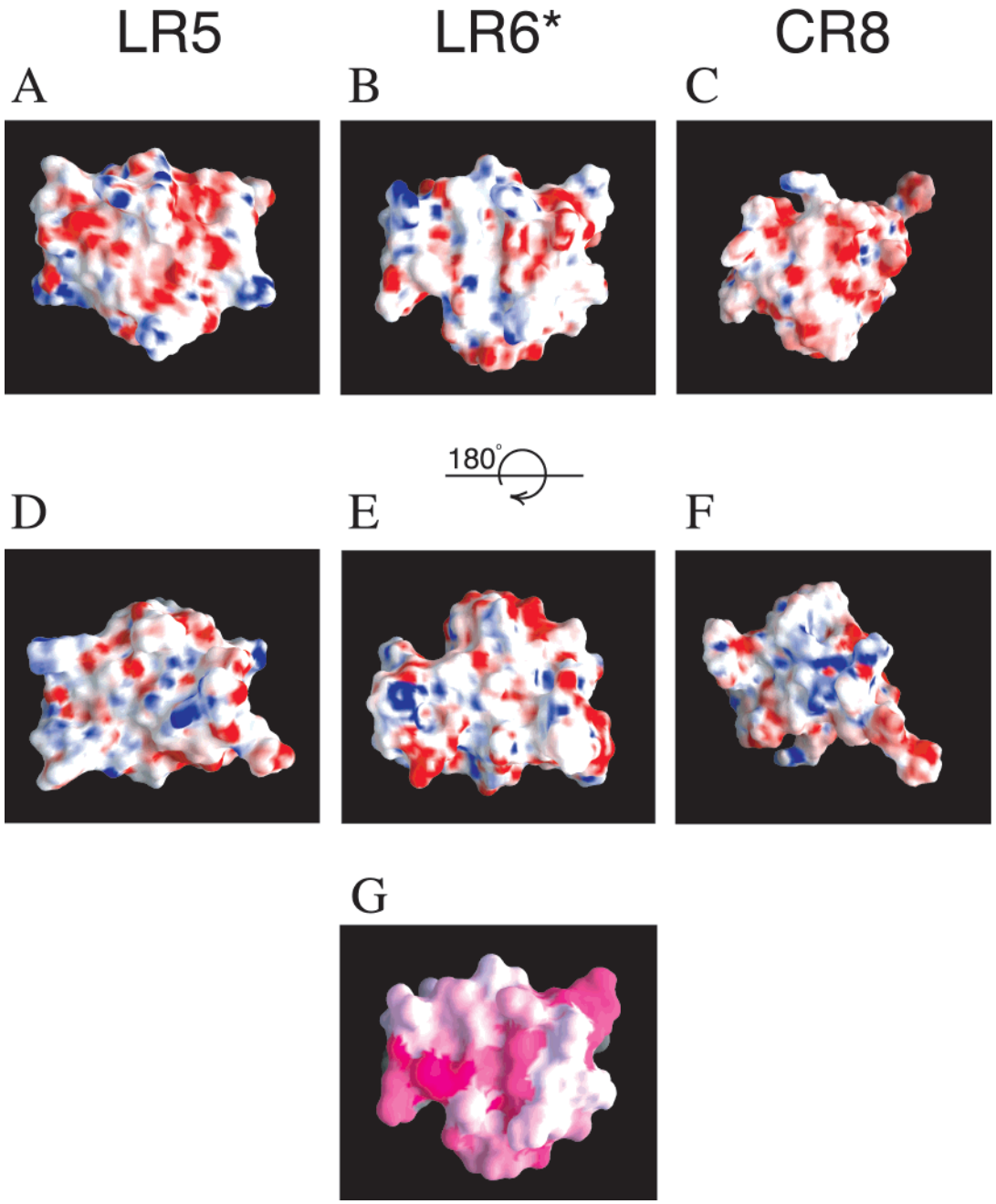

FIGURE 5: (A−F) Molecular surfaces of the structures of LR5 (A and D), LR6* (B and E), and CR8 (C and F). The surface color is defined by electrostatic potential (red, most negative, to blue, most positive), calculated as described in Materials and Methods. In panels A−C, modules are illustrated in an orientation similar to that in Figure 3. In panels D−F, the modules are shown rotated 180° around the horizontal axis. (G) Sequence variability of LDL-A modules, extracted from the HSSP database, mapped to the surface of LR6*. A sliding scale from dark pink (conserved) to white (highly variable) is shown. This figure was prepared with the program GRASP (*26*).

N-terminal lobe of LDL-A modules (*10*). This notion is consistent with the observation that the amide protons of these residues are the ones most highly protected against hydrogen exchange within LR6*.

LDL-A Modules Contain a Conserved Surface Patch of Negative Electrostatic Potential. It is well established that a positively charged region on the surface of the fourth helix (*35*) of the N-terminal domain of apolipoprotein E (ApoE) is required for binding of ApoE-containing particles to the LDLR (*36*−*38*). The presence of a conserved acidic motif within LDL-A (ligand-binding) modules thus led to the proposal that electrostatic complementarity between receptor and ligand is primarily responsible for lipoprotein binding by the LDLR (*1*, *35*, *39*, *40*). This suggestion was recently challenged on the basis of the crystal structure of LR5, because the conserved acidic motif appears primarily to play a structural role in coordinating a bound calcium ion (*10*).

However, among the three LDL-A modules for which the structure has been determined with explicit consideration of the calcium-binding site, the most striking similarity is the fact that there is a region of negative electrostatic surface potential that surrounds this coordinated ion (Figure 5A−F). Although the presence of the tightly bound calcium ion attenuates the negative surface potential resulting from the conserved acidic motif, it appears that the presence of the bound ion serves to localize and arrange the negatively

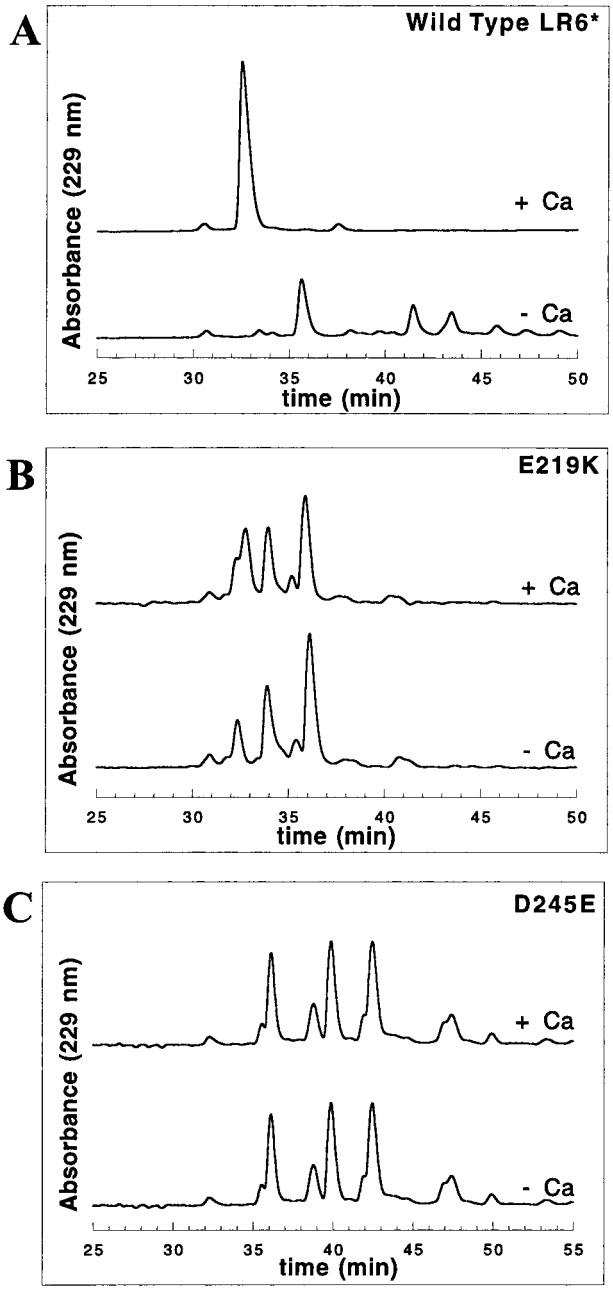

FIGURE 6: Folding of ligand-binding module 6 (LR6*) of the LDL receptor. (A) When LR6* is folded in the presence of calcium (see Materials and Methods), a single predominant disulfide isomer is formed (top). When calcium is omitted from the folding buffer, a distribution of non-native isomers results (bottom). The E219K (B) and D245E (C) point mutants of LR6*, both found in FH patients, do not fold to a single disulfide isomer, even in the presence of calcium. In each case, the distribution of isomers is similar in the presence (top) or absence (bottom) of calcium.

charged side chains that might otherwise repel one another.

The high degree of conservation of the residues that comprise this negatively charged surface strongly suggests that these electrostatic properties are an intrinsic feature of most, if not all, LDL-A modules (Figure 5G). As a result, the general preference of LDL-A modules for highly basic ligands may indeed result from an electrostatic component in ligand binding.

In contrast, the basis for ligand-binding specificity may be derived from the unique surface features of individual modules or strings of modules, rather than from the highly conserved calcium-binding site. For example, the surface contours of individual modules are markedly different (Figure 5A−F). In addition, it is likely that binding sites for many specific ligands consist of composite surfaces specified by two or more adjacent modules.

Analysis of LR6 FH Mutations. Two variants of LR6*, identified originally as FH mutations, have been investigated for their ability to form native disulfide bonds under conditions that permit disulfide exchange. The first, E219K, lies near the amino-terminal end of LR6* at a position that is not conserved among LDL-A modules, whereas the second, D245E, alters one of the aspartate side chains that directly coordinate the bound calcium ion. After equilibration at physiologic calcium concentrations, neither E219K nor D245E folds to a unique disulfide isomer, indicating that these FH mutations give rise to a protein-folding defect (Figure 6).

From the LR6* structure, it appears that the E219K mutation would disrupt a hydrogen bond that normally exists between the side chain carboxylate of Glu 219 and the backbone amide proton of Arg 216. The surprising observation that E219K interferes with calcium-dependent folding of LR6* argues for cooperativity between the N- and C-terminal lobes of the module in specifying the LDL-A fold.

The observation that the D245E mutation prevents proper folding of LR6* strengthens previous claims (*11*, *27*) that calcium coordination is crucial for proper folding of all LDL-A modules, and illustrates that even a subtle alteration of this calcium-binding site is sufficient to disrupt folding of LDL-A modules. Thus, it is likely that the D235G mutation, which eliminates entirely one of the side chain carboxylates of the calcium-binding site, will also prevent proper folding of LR6*.

Other FH mutations within LR6 likely to prevent proper module folding include those that alter five of the six cysteine residues that form the three stabilizing disulfide bonds. Similarly, the Q233P mutation, which immediately precedes the conserved cysteine and aspartate at positions 234 and 235, respectively, is likely to disrupt the local backbone geometry enough to interfere with proper folding. The mechanism by which the S223Y and R232W mutations might cause FH, however, remains unclear, especially given that a tryptophan residue occupies the position corresponding to R232 in four of the seven LDL-A modules of the LDLR.

ACKNOWLEDGMENT

We thank Susan Pochapsky of the MIT/Harvard Center for Magnetic Resonance and Gregory Heffron of Gerhard Wagner's laboratory for technical assistance with NMR spectrometers, Dr. Gerhard Wagner and the Center for Structural Biology for the use of NMR spectrometers at Harvard Medical School, and the MIT/Harvard Center for Magnetic Resonance for the use of spectrometers housed at MIT. We are indebted to Jean-Paul Font for purifying the E219K and D245E mutants of LR6*, and for performing disulfide bond rearrangements.

SUPPORTING INFORMATION AVAILABLE

Table of statistics for coordinate precision (i.e., rmsd to mean for backbone and heavy atoms) for structures of LR6* when H-bonds, calcium restraints, or both H-bonds and

calcium restraints are omitted from the structure calculations and a figure depicting the mean structure from each of these ensembles and comparing these calculated structures with the final reported mean structure (from Figure 3 of the main body of the paper) in an overlay plot of the polypeptide backbone. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Brown, M. S., and Goldstein, J. L. (1986) *Science 232*, 34−47.
2. Gliemann, J. (1998) *Biol. Chem. Hoppe-Seyler 379*, 951−964.
3. Herz, J., Hamann, U., Rogne, S., Myklebost, O., Gausepohl, H., and Stanley, K. K. (1988) *EMBO J. 7*, 4119−4127.
4. Saito, A., Pietromonaco, S., Loo, A. K., and Farquhar, M. G. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9725−9729.
5. Kim, D. H., Iijima, H., Goto, K., Sakai, J., Ishii, H., Kim, H. J., Suzuki, H., Kondo, H., Saeki, S., and Yamamoto, T. (1996) *J. Biol. Chem. 271*, 8373−8380.
6. Takahashi, S., Kawarabayasi, Y., Nakai, T., Sakai, J., and Yamamoto, T. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9252−9256.
7. Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L., and Russell, D. W. (1984) *Cell 39*, 27−38.
8. Russell, D. W., Brown, M. S., and Goldstein, J. L. (1989) *J. Biol. Chem. 264*, 21682−21688.
9. Goldstein, J. L., Hobbs, H. H., and Brown, M. S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. S., Beaudet, A. L., Sly, W. S., and Valle, D., Eds.) pp 1981−2030, McGraw-Hill Inc., New York.
10. Fass, D., Blacklow, S., Kim, P. S., and Berger, J. M. (1997) *Nature 388*, 691−693.
11. Blacklow, S. C., and Kim, P. S. (1996) *Nat. Struct. Biol. 3*, 758−762.
12. North, C. L., and Blacklow, S. C. (1999) *Biochemistry 38*, 3926−3935.
13. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol. 154*, 367−382.
14. Bartels, C., Xia, T., Güntert, P., Billeter, M., and Wüthrich, K. (1995) *J. Biomol. NMR 5*, 1−10.
15. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185−189.
16. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
17. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) *J. Am. Chem. Soc. 111*, 1515−1517.
18. Marion, D., and Bax, A. (1988) *J. Magn. Reson. 80*, 528−533.
19. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins 17*, 75−86.
20. Güntert, P., Berndt, K. D., and Wüthrich, K. (1993) *J. Biomol. NMR 3*, 601−606.
21. Kay, L. E., Brooks, B., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) *J. Am. Chem. Soc. 111*, 5488−5490.
22. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
23. Guntert, P., Braun, W., and Wuthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
24. Brünger, A. T. (1992) *X-PLOR Version 3.1. A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
25. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
26. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−296.
27. Atkins, A. R., Brereton, I. M., Kroon, P. A., Lee, H. T., and Smith, R. (1998) *Biochemistry 37*, 1662−1670.
28. Huang, W., Dolmer, K., and Gettins, P. G. (1999) *J. Biol. Chem. 274*, 14130−14136.
29. Djordjevic, J. T., Bieri, S., Smith, R., and Kroon, P. A. (1996) *Eur. J. Biochem. 239*, 214−219.
30. Marazziti, D., Eggersten, G., Fey, G. H., and Stanley, K. K. (1988) *Biochemistry 27*, 6529−6534.
31. Tensen, C. P., Van Kesteren, E. R., Planta, R. J., Cox, K. J., Burke, J. F., van Heerikhuizen, H., and Vreugdenhil, E. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4816−4820.
32. Bates, P., Young, J. A. T., and Varmus, H. E. (1993) *Cell 74*, 1043−1051.
33. Daly, N. L., Scanlon, M. J., Djordjevic, J. T., Kroon, P., and Smith, R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6334−6338.
34. Daly, N. L., Djordevic, J. T., Kroon, P. A., and Smith, R. (1995) *Biochemistry 34*, 14474−14481.
35. Wilson, C., Wardell, M. R., Weisgraber, K. H., Mahley, R. W., and Agard, D. A. (1991) *Science 252*, 1817−1822.
36. Innerarity, T. L., Friedlander, E. J., Rall, S. C., Jr., Weisgraber, K. H., and Mahley, R. W. (1983) *J. Biol. Chem. 258*, 12341−12347.
37. Lalazar, A., Weisgraber, K. H., Rall, S. C., Jr., Giladi, H., Innerarity, T. L., Levanon, A. Z., Boyles, J. K., Amit, B., Gorecki, M., Mahley, R. W., et al. (1988) *J. Biol. Chem. 263*, 3542−3545.
38. Weisgraber, K. H., Innerarity, T. L., Harder, K. J., Mahley, R. W., Milne, R. W., Marcel, Y. L., and Sparrow, J. T. (1983) *J. Biol. Chem. 258*, 12348−12354.
39. Sudhof, T. C., Goldstein, J. L., Brown, M. S., and Russell, D. W. (1985) *Science 228*, 815−822.
40. Mahley, R. W. (1988) *Science 240*, 622−630.
41. Springer, T. A. (1998) *J. Mol. Biol. 283*, 837−862.
42. Wilson, D. J., Gahan, M., Haddad, L., Heath, K., Whittall, R. A., Williams, R. R., Humphries, S. E., and Day, I. N. (1998) *Am. J. Cardiol. 81*, 1509−1511.
43. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
44. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 29−32.

BI992087A