Three-Dimensional Structure of the LDL Receptor-Binding Domain of Human Apolipoprotein E

CHARLES WILSON, MARK R. WARDELL, KARL H. WEISGRABER, ROBERT W. MAHLEY, DAVID A. AGARD*

Human apolipoprotein E, a blood plasma protein, mediates the transport and uptake of cholesterol and lipid by way of its high affinity interaction with different cellular receptors, including the low-density lipoprotein (LDL) receptor. The three-dimensional structure of the LDL receptor-binding domain of apoE has been determined at 2.5 angstrom resolution by x-ray crystallography. The protein forms an unusually elongated (65 angstroms) four-helix bundle, with the helices apparently stabilized by a tightly packed hydrophobic core that includes leucine zipper–type interactions and by numerous salt bridges on the mostly charged surface. Basic amino acids important for LDL receptor binding are clustered into a surface patch on one long helix. This structure provides the basis for understanding the behavior of naturally occurring mutants that can lead to atherosclerosis.

LIPIDS, INCLUDING TRIGLYCERIDES, PHOSPHOLIPIDS, AND cholesterol, are sparingly soluble in aqueous solution and are thus transported through the body in the form of lipoprotein particles. Apolipoproteins, the protein components of these particles, take part in stabilizing lipoproteins and directing their metabolism. Two apolipoproteins, apolipoprotein E (apoE) and apolipoprotein B (apoB), bind with high affinity to cell surface receptors, including the low-density lipoprotein receptor, and thereby mediate the cellular uptake of most lipoproteins, namely very low density lipoprotein (VLDL), low-density lipoprotein (LDL), and high-density lipoprotein (HDL) (*1*). Because plasma cholesterol concentrations and metabolism are unequivocally linked to the development of atherosclerosis and risk of coronary artery disease, it is of fundamental importance to characterize the interaction of apoE and apoB with the LDL receptor at the molecular level. As a first step, we have pursued high-resolution crystallographic studies of apoE.

ApoE is a 299-residue protein (relative molecular mass, 34,200 daltons) that appears to be made up of two independently folded domains (*2*). Digestion with thrombin produces a 22-kD fragment (residues 1 to 191) corresponding to the NH_2-terminal domain and a 10-kD fragment (residues 216 to 299) corresponding to the COOH-terminal domain (Fig. 1) (*2*). Extensive characterization of apoE has revealed that the structural domains also define its functional domains. While the NH_2-terminal region is responsible for the binding of apoE to the LDL receptor (*3*) and binds lipids only weakly (*4, 5*), the COOH-terminal region mediates the binding of apoE to the surface of lipoproteins (*4, 5*), but does not bind to the LDL receptor (*3*).

Although the NH_2-terminal domain of apoE is related to other members of the apolipoprotein gene family (apoAI, apoAIV, apoCI, apoCII, and apoCIII), it is fundamentally different in some respects. All these proteins contain internal sequence repeats predicted to form amphipathic alpha helices (*6, 7*), and circular dichroism spectroscopy has shown that they all have a high helical content (*8*). However, whereas most apolipoproteins are only marginally stable and form aggregates in solution in the absence of lipid, the NH_2-terminal domain of apoE exists as a monomer in solution at high concentrations, and its free energy of stabilization ($\Delta G_{unfold} \approx 10$ kcal/mol) is typical of other globular soluble pro-

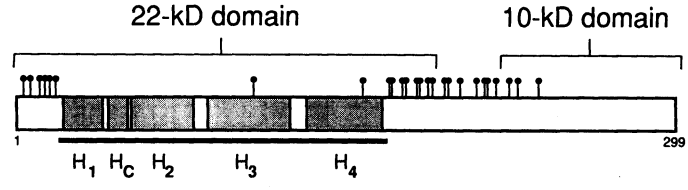


Fig. 1. Domain structure of apoE. Residues 1 to 191 define the NH_2-terminal domain of apoE, known to be important for LDL receptor binding. Solid circles indicate proteolytically sensitive residues [as determined by digestion with thrombin, elastase, trypsin, chymotrypsin, *Staphylococcus aureus V8*, and subtilisin (*2*)]. The solid bar indicates the region that is included in the refined crystal structure of the 22-kD fragment (residues 24 to 166). The helices (shaded areas) making up the four-helix bundle are numbered as follows: H1, residues 24 to 42; H2, 54 to 81; H3, 87 to 122; and H4, 130 to 164. A connecting helix (Hc, residues 44 to 53) joins helices 1 and 2. Secondary structure assignments were made by the method of Kundrot and Richards (*43*). The complete apoE sequence includes three weakly conserved 11-residue repeats (29–39, 40–50, 51–61) and eight 22-residue repeats (62–83, 84–105, 106–127, 128–149, 150–171, 172–193, 194–215, 216–237) (*6, 7*).

C. Wilson and D. A. Agard are associated with the Howard Hughes Medical Institute, the Graduate Group in Biophysics, and the Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143-0448. M. R. Wardell, K. H. Weisgraber, and R. W. Mahley are at the Gladstone Foundation Laboratories for Cardiovascular Disease, Cardiovascular Research Institute, Department of Pathology, University of California, San Francisco, 94140. The present address of C. Wilson is the Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114 and that of M. R. Wardell is the Department of Hematology, University of Cambridge, Cambridge, U.K. CB2 2QH.

*To whom correspondence should be addressed.

teins (2). By contrast, the COOH-terminal domain of apoE exhibits properties more typical of the other soluble apolipoproteins; it forms a tetramer in solution, and displays a relatively low free energy of stabilization ($\Delta G_{unfold} \approx 3$ to 4 kcal/mol).

The properties of the NH_2-terminal domain of apoE made it an attractive candidate for crystallographic analysis, and indeed it proved possible to grow high-quality crystals (9). In contrast, the lack of suitable crystals for any other vertebrate apolipoprotein has made it impossible to perform high-resolution structural studies of these apolipoproteins. We report here the atomic structure of the 22-kD NH_2-terminal thrombolytic fragment of human apoE, based on the interpretation of an electron density map at 2.5 Å.

Structure determination. Possibly as a result of slight changes in the protein purification procedure, we found it difficult to crystallize the NH_2-terminal domain when we followed our previously reported procedure (9) and therefore we searched for new crystallization conditions. Crystals suitable for x-ray diffraction studies were obtained with the 22-kD thrombolytic fragment of human apoE3 by the hanging drop method; with the use of protein at 10 mg/ml [prepared as described in (9), except for the addition of high-performance liquid chromatography (HPLC)], 15 percent PEG 400 (BDH), with 20 mM sodium acetate–acetic acid buffer, 0.2 percent β-*n*-octylglucopyranoside (Calbiochem), and 0.1 percent β-mercaptoethanol (final concentrations). Crystallization was induced by raising the pH of the crystallization reservoir from 4.5 to 5.3 (fully grown crystals were stable up to pH 7.2). The detergent β-*n*-octylglucopyranoside was not required for crystal growth but was helpful in reducing twinning. Using these conditions, we were able to routinely grow small perfect prisms. Precession photography showed that, although the morphology of these crystals differed significantly from those reported (9), they belonged to the same space group ($P2_12_12_1$) and had approximately the same unit cell dimensions (a = 41.3 Å, b = 54.5 Å, c = 87.0 Å, $\alpha = \beta = \gamma \equiv 90°$).

Data were collected with a RIGAKU AFC5R diffractometer equipped with a graphite monochromator and a 600-mm detector arm. Early experiments indicated that the crystals were extremely susceptible to radiation decay. Crystals were rapidly frozen in a stream of boiling liquid nitrogen, and data were collected at −160°C with a low-temperature apparatus (MSC; College Station, Texas). Freezing resulted in a slight shrinkage of the unit cell (a = 40.7 Å, b = 54.0 Å, c = 85.4 Å), as has been seen in most other low-temperature protein crystallography experiments (10). Freezing reduced radiation decay, making it possible to collect data for up to 2 weeks on a single crystal (Table 1).

A Patterson map was calculated from the native data to 4 Å resolution. The map contained a number of strong peaks (corresponding to overlapping interatomic vectors), which suggested the presence of a four-helix bundle. Standard molecular replacement techniques (11), with an idealized four-helix bundle as a search model indicated two possible configurations. Although rigid body refinement of either a four-helix bundle or a six-helix bundle (made up by two overlapping four-helix bundles) caused a reduction in the overall R factor (10 to 4 Å data) to 44 to 48 percent (12), we were unable to identify portions of the molecule missing from the search model. This was most likely due to our incorrect assumption that the helices were 16 to 18 residues long.

We screened approximately 40 different ionic heavy-atom compounds by diffractometry, none of which yielded a suitable isomorphous derivative. Although apoE3 contains a single free cysteine (residue 112), none of the reactive mercurials tested were successful. However, equilibrating crystals with dimethyl mercury (a small, nonreactive, nonionic compound) did produce a useful derivative. Heavy atom refinement with the TCTREF program (13) indicated two high-occupancy sites. Two long helices were easily identified in a map calculated with the heavy atom phases. Density modification (by solvent flattening) was carried out with the programs of Wang (14) and, while noise in the map was significantly reduced, the electron density remained too ambiguous to allow modeling of the structure (Fig. 2A). To improve the phase estimates, we collected accurate anomalous data on the mercury derivative for those reflections predicted to have the strongest anomalous scattering signal (Table 1). Combined phase estimates together with density modification resulted in an easily interpretable map (Fig. 2B).

A model containing residues 23 to 164 was constructed, and three cycles of X-PLOR (15) refinement and manual rebuilding reduced the crystallographic R factor for all data in the 8 to 2.5 Å range to 23.0 percent. The heavy atoms were found to occupy a hydrophobic pocket on the surface formed by Trp^{34} and three neighboring leucines. The addition of 76 bound solvent molecules and residues 165–166 further reduced the R factor to 17.7 percent. Individual isotropic B factors were refined for all atoms in the model. The electron density was calculated with phases based on the refined model and $2F_o$-F_c coefficients (Fig. 2C). The root-mean-square (rms) deviation of bond lengths (0.017 Å) and bond angles (3.2°) for the refined model are typical of well-determined protein structures. All backbone dihedral angles (ϕ-Ψ pairs) fall within allowed regions of the Ramachandran map, except for glycines at 127 and

Table 1. Statistics for crystallographic data and refinement. Data were collected at liquid nitrogen temperatures with a Rigaku AFC5R diffractometer (rotating anode generator operated at 180 mA, 50 kV). The native crystals belong to space group $P2_12_12_1$ and showed unit cell dimensions a = 40.65 Å, b = 53.96 Å, c = 85.43 Å. Native data represent the merged observations from three crystals. The isomorphous derivative data were obtained by diffusing dimethyl mercury into fully grown crystals for 2 weeks. Isomorphous data were collected from two crystals and anomalous data were collected from a single crystal (statistics for anomalous data are shown in parentheses).

Items	Native	$(CH_3)_2Hg$ derivative	
Resolution (Å)	2.5	2.5	(2.8)
Diffraction data			
Observations (N)	17,129	9,957	(3,317)
Unique reflections (N)	6,899	6,880	(1,499)
Completeness (percent)	100	100	(31)
$<I>/<\sigma_I>$	43.48	30.42	(51.58)
R_{merge}*	0.054		
Phasing statistics			
Heavy atom sites		2	
Isomorphous differences		6,474	
Anomolous differences			(1,244)
rms F_H/residual†		1.55	(1.57)
R_{Cullis}		0.55	(0.50)
rms R_{anom}§			(0.48)
<Figure of merit>		0.41	(0.53)
<Figure of merit following solvent flattening>		0.56	
X-PLOR refinement statistics			
Protein atoms (non-hydrogens) (N)	1,172		
Water molecules (N)	76		
R_{cryst} (overall)	0.177		
8.50 Å to 3.85 Å	0.149		
3.85 Å to 3.12 Å	0.171		
3.12 Å to 2.74 Å	0.204		
2.74 Å to 2.50 Å	0.223		
rms deviations from ideality			
bond lengths (Å)	0.017		
bond angles (°)	3.2		

*$R_{merge} = (\Sigma\Sigma|I_j(h) - <I(h)>|)/(\Sigma\Sigma I_j(h))$ (summations done over all reflections in all crystals.) †Root-mean-square F_H/residual = phasing power where F_H = rms heavy atom structure factor amplitude (calculated) and residual = rms lack of closure error. ‡$R_{Cullis} = (\Sigma|FPH_{calc} - FPH_{obs}|)/(\Sigma|FPH_{obs} - FP|)$ (for centric reflections only, FPH and FP are structure factors for the derivative and native protein, respectively). §Root-mean-square R_{anom} = rms phase-averaged anomalous residual divided by the rms observed anomalous difference. $R_{cryst} = (\Sigma|F_{obs} - F_{calc}|)/(\Sigma|F_{obs}|)$.

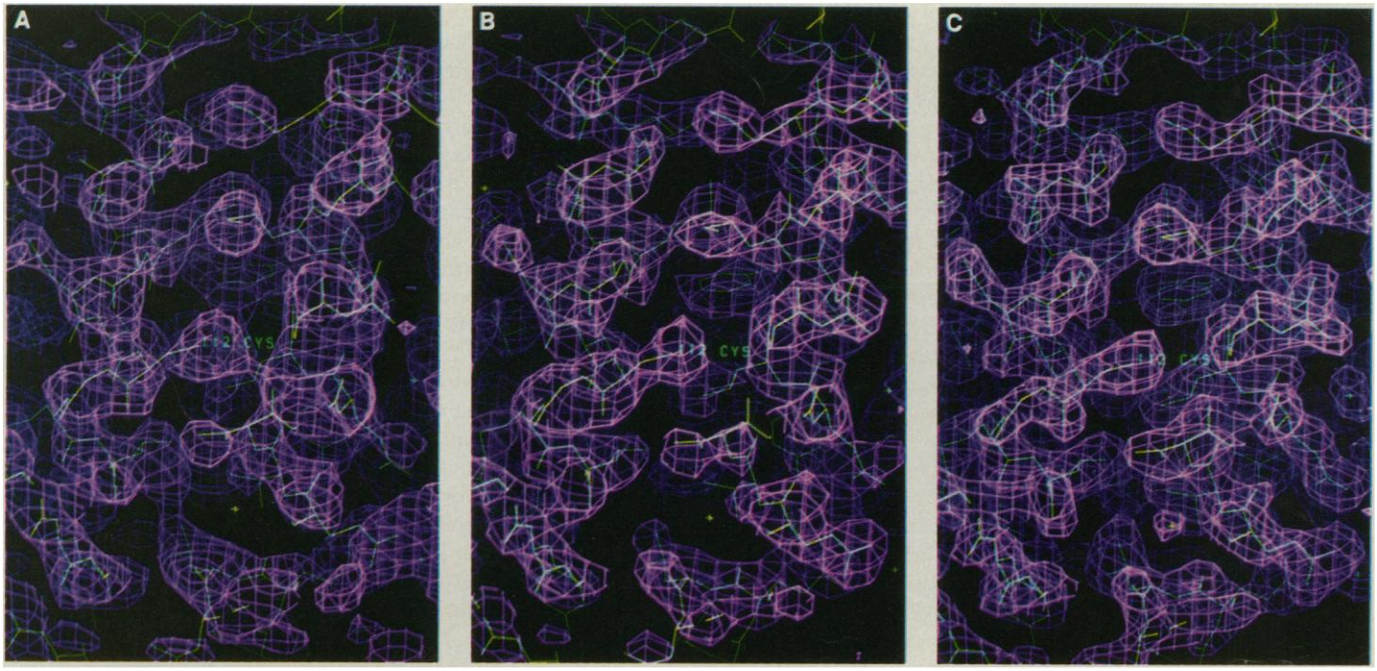


Fig. 2. Electron density calculated at various stages of phase refinement. The region surrounding Cys^{112}, site of the mutation in the apoE4 isoform is shown. (**A**) Electron density map calculated with the use of SIR phases from the dimethyl mercury derivative, after extensive solvent flattening. (**B**) Map calculated with SIR plus anomalous phases, also after solvent flattening. (**C**) $2F_0$-F_c map calculated with the final α_c phases.

165. Of all side chains, 92 percent can be classified as rotamers in the Ponder and Richards library (*16*). Most of the non-rotamer side chains are glutamate and methionine residues, which are poorly sampled in the Ponder and Richards data set.

Our final model lacks amino acid residues 1 to 22 and 167 to 191. At both ends of the modeled region, density is well determined but quickly disappears outside the model. Electrophoresis of the crystallized protein demonstrated that the complete 22-kD fragment was present; that is, the missing regions were not absent as a result of proteolysis. We believe residues 1 to 22 and 167 to 191 are disordered in the crystal and are not missing because of systematic phasing errors: (i) Digestion of apoE with a battery of proteases shows that amino acids 1 to 20 and 165 to 191 are susceptible to proteolysis and thus likely to be unstructured (Fig. 1) (*2*). (ii) The original SIR-SAS map, which is unbiased by the model, shows that there is no significant density remaining outside the modeled region. (iii) A difference map calculated between the bacterially expressed human protein and the native plasma-derived human protein fails to show a peak corresponding to the additional NH_2-terminal methionine present in the bacterial protein, even at low resolution. (iv) The *R* factor is unlikely to drop to less than 18 percent with excellent stereochemistry if the missing residues contribute significantly to the scattering.

ApoE: An unusually elongated four-helix bundle. The NH_2-terminal domain of apoE contains five helices comprising more than 80 percent of the modeled residues (Fig. 3). Four of the helices, containing 19, 28, 36, and 35 amino acids each, are arranged to form a 2 by 2 bundle. Four-helix bundles are the most common tertiary fold in α-helical proteins and are found in at least 18 other protein crystal structures (*17*). The average length for bundle helices in these structures is 18 residues, roughly half the helix length found for apoE. Comparison to the core of myohemerythrin (*18*), a typical four-helix bundle, indicates that the spacing between helices and the interhelical angles are similar for the two proteins.

Each helix in the apoE bundle lies antiparallel to the helices adjacent to it [this up-down topology is found in all other bundles except cytochrome $P\text{-}450_{cam}$ (*17*)]. The connection between the first two helices of the bundle (residues 24 to 42 and 54 to 81) is a short helix (residues 44 to 53). The turn between helix 2 and helix 3 of the bundle (residues 82 to 86) is poorly defined in the electron density map and in the refined model. The average crystallographic B factor for atoms in this loop is more than 60 $Å^2$; completely omitting them causes an insignificant increase in the overall *R* factor. The residues in this loop have been included in the final model for completeness, but their coordinates may be somewhat in error. Helix 3 (residues 87 to 122) is kinked at Gly^{105} but maintains main chain hydrogen bonding through this region. Helix 4 (residues 130 to 164), containing residues known to be important in LDL receptor binding, is well ordered throughout.

Internal sequence repeats and apolipoprotein evolution. As mentioned earlier, ApoE shares a gene structure common to most other apolipoproteins (apoAI, apoAIV, apoCI, apoCII, and apoCIII) (*6, 7, 19*). In all of these proteins, the coding region of the mature peptide is split between exons three and four. Exon 3 includes three consecutive 11–amino acid repeats while exon 4 contains 1 to 12 copies of a 22–amino acid repeat (*6, 7*). The protein sequence identity between the apoE repeats is relatively low compared with that in other apolipoproteins and is barely detectable in the nucleotide coding sequence. Figure 3 shows a clear structural role for these repeats as the basis for the long helices in the bundle. Helix 1 and the connecting helix are defined by the first two 11-residue repeats of exon 3. Helix 2 contains the third 11-residue repeat and one 22-residue repeat, while helices 3 and 4 are each made up of two 22-residue repeats. The apparently disordered COOH-terminal residues (167 to 191) correspond to parts of the fifth and sixth 22-residue repeats. Multiple copies of short internal repeats have been detected in several different protein sequences [for example, the insect protein apolipophorin III (*20*), the leucine-rich α_2-glycoprotein protein of human serum (*21*), and the tau and MAP2 microtubule-binding proteins (*22, 23*)]. The crystal structure of apoE provides clear evidence that repeats such as these can function as structural building blocks within a single domain.

The boundaries between internal repeats appear to have two alternative functions in the folded structure. At residues 40, 51, 84, and 128, the junctions form turns, placing neighboring repeats into separate helices. At residues 62, 106, and 150, the junctions adopt a helical conformation and maintain the two adjacent repeats in a single helix. The helix built up by the third and fourth 22–amino acid repeats (helix 3) is strongly kinked at the repeat junction (residues 105 to 106). None of the turns in the structure correspond to the middle of a repeat. This suggests that, while the ends of the repeats are not limited to one specific conformation, the cores of the repeats are inherently helical. In a simple model for the structure and evolution of apolipoproteins, the internal repeats correspond to stable helices that can pack well with their helical axes antiparallel to one another. Whether adjacent sequence repeats exist in the same or in separate helices is likely to be determined by the intervening junction sequences. An understanding of the sequence-dependence of stability is needed to predict the topology of other apolipoproteins.

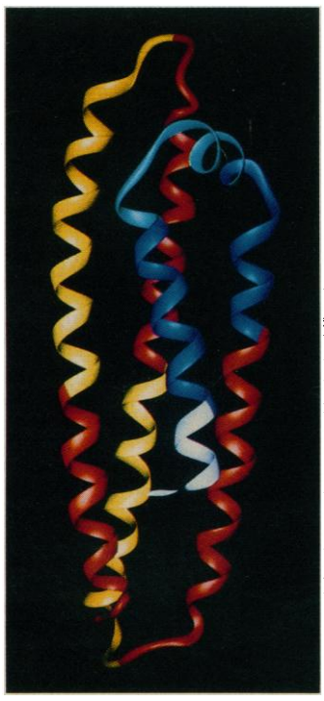
As predicted from analysis of apolipoprotein sequences (*24*), the helices of apoE are strongly amphipathic. The four long helices of the bundle are arranged such that their hydrophobic residues are completely sequestered inside the protein while their hydrophilic faces are solvent exposed (Fig. 4). Leucine side chains occurring about every seven residues appear to stabilize the interfaces between helix 1 and helix 4 and between helix 2 and helix 3 (Fig. 5). The organization of leucine interactions on adjacent antiparallel helices is similar to the leucine zipper model of Landschulz *et al.* (*25*), predicted for the dimerization domain of C/EBP-type transcription factors. Although subsequent experimental work has shown that transcription factor dimerization is probably not based on antiparallel helices (*26*), the crystal structure of apoE indicates that the original leucine zipper model is an energetically feasible way of stabilizing pairs of helices.

More than a third of the NH_2-terminal domain residues are charged. The crystal structure includes atomic coordinates for 24 acidic and 24 basic residues. These amino acids almost completely cover the surface of the bundle (Fig. 4), and most participate in either intramolecular or intermolecular salt bridges in the crystal structure. Of the intramolecular salt bridges, eleven are formed by pairs of amino acids lying in the same helix. Surprisingly, more than half of these salt bridges are oriented to interact unfavorably with the macrodipole of the helix to which they belong. Seven salt bridges are formed between pairs of helices in the bundle. These strong electrostatic interactions may help to bind the helices together and further stabilize the folded structure. The combination of tight hydrophobic packing and numerous electrostatic interactions would seem to account for the high free energy of stabilization for this domain, setting it apart from most other apolipoproteins.

An electrostatic potential map was calculated for the 22-kD fragment with the DELPHI program (Fig. 6) (*27*). The DELPHI algorithm solves the Poisson-Boltzmann equation with the use of the formal charge distribution indicated by the crystal structure. Because most charged residues are paired to form salt bridges, the net electrostatic potential is close to zero for most of the region surrounding the protein. The only significant feature of this map is a large region of positive potential encompassing the NH_2-terminal half of helix four (residues 136 to 150). The possible role of this electrostatic feature in LDL receptor binding is discussed below.

Basis for LDL receptor binding and lipid binding. One of apoE's major functions is in mediating the cellular uptake of lipoprotein particles (*1*). Lipoprotein-associated apoE binds with high affinity to LDL receptors on the surface of target cells. Upon apolipoprotein binding, these receptors cluster in clathrin-coated membrane pits, which subsequently pinch off to internalize the lipoprotein particle. This process eventually results in the degradation of the lipoprotein particle and the release of its lipids for cellular use. ApoE also functions in the uptake of chylomicron remnants by liver cells via an apoE-specific receptor—possibly LRP, the LDL receptor–related protein (*28*).

Low-density lipoprotein receptor binding is localized in the NH_2-terminal domain of apoE (*3*). Genetic analysis of point mutations in apoE, in combination with site-directed mutagenesis and antibody competition studies, have shown that residues 136 to 150 are necessary for this function (*3, 29*). Electron density for all of these residues, found to lie on helix 4 of the bundle, is well determined. The NH_2-terminal portion of helix 4 is unusually rich



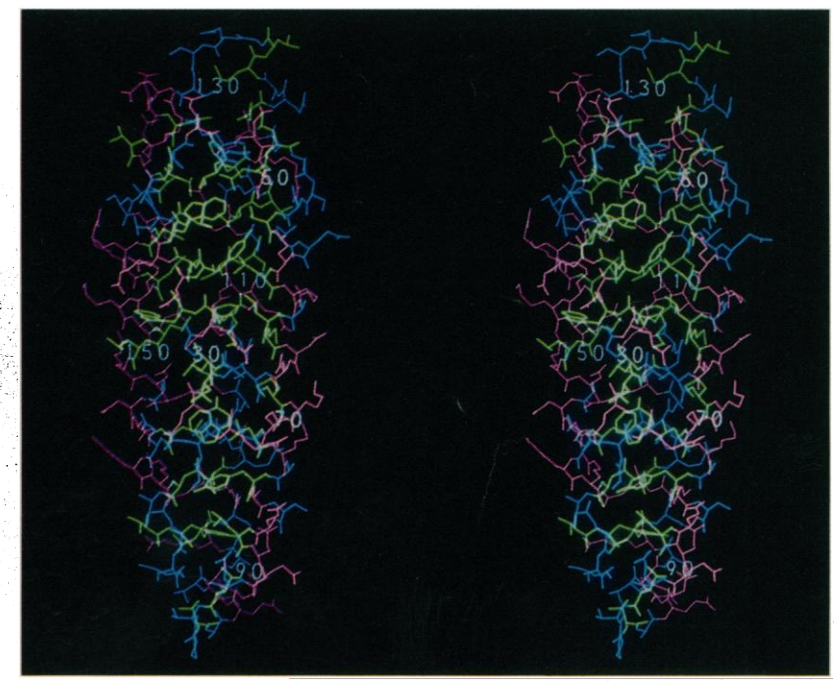


Fig. 3 (**left**). Ribbon diagram of the NH_2-terminal domain based on the crystal structure. Internal sequence repeats are used to color code the α helices. The three 11-residue repeats are colored blue and cyan, while the five 22-residue repeats are colored red and yellow. **Fig. 4** (**right**). Stereo view of the refined atomic model of the LDL receptor-binding domain. Hydrophobic residues (green) line the buried faces of the bundle helices and form a well-packed core. Charged amino acids (red) cover most of the surface.

in basic residues, including Arg^{134}, Arg^{136}, His^{140}, Arg^{142}, Lys^{143}, Arg^{145}, Lys^{146}, Arg^{147}, and Arg^{150}. Because most of these amino acids are solvent-exposed and not involved in intramolecular salt bridges, their excess positive charge combines to produce a large region of positive electrostatic potential extending 15 Å out of the protein (Fig. 6). Naturally occurring variants are known for positions 136, 142, 145, and 146 in which the basic amino acid is substituted by a neutral (*3, 25*) or in one case by an acidic amino acid (*30*). All of these variants display defective binding to the LDL receptor (20 to 40 percent of normal binding). Site-directed mutagenesis of other basic residues in the helix 4 cluster, Lys^{143} or Arg^{150}, also reduces receptor binding to 10 to 50 percent of normal levels (*31*).

The fact that no single substitution results in complete abolition of binding activity suggests that there are multiple interactions between apoE and the receptor and that the basic residues may cooperate in the binding function. That these basic residues are not involved in salt bridges within the structure supports the hypothesis that they are free to interact with the receptor.

The LDL receptor recognizes both apoE and apoB, with comparable affinity (*32*). While there is essentially no sequence similarity between these two ligands, apoB contains a region rich in basic residues that has a pattern of charged and neutral amino acids similar to that in the receptor-binding helix of apoE. Several studies have indicated that this region may play a role in receptor binding (*33*). It is possible that a positively charged helix is used by both proteins to specify LDL receptor binding. Because of apoB's insolubility and large size (more than 500 kD), crystallographic solution of the structure of this medically relevant protein may not be available in the near future. Therefore, the structure of apoE's receptor-binding helix may provide a useful starting point for modeling the interaction of this other important ligand with the LDL receptor.

Although the binding of apoE to lipoprotein particles appears to be mediated by the COOH-terminal domain (residues 216 to 299) (*4, 5*), the NH_2-terminal domain does associate in vitro with phospholipid to form discoidal particles (*3*). In the lipid-bound form apoE is able to bind with very high affinity to the LDL receptor ($K_d \approx 1.2 \times 10^{-10}$ M) (*34*); recent work suggests that the binding affinity of lipid-free protein is about 500 times lower than that of the lipid-complexed protein (*35*). Several studies had suggested that helices in the NH_2-terminal domain (especially helix 4) may insert into phospholipid membranes. In support of this, short synthetic peptides corresponding to helix 4 have been shown to have lipid binding activity (*36*). By contrast, the crystal structure does not indicate a special lipid binding role for helix 4 or any of the other helices. Hydrophobic amino acids in the 136 to 150 region are all buried and well packed, interacting with the hydrophobic residues of helices 1 and 3. In addition, the α-helical hydrophobic moments for all of the core helices fall in the same range as those for the helices in other four-helix bundle proteins, none of which have significant lipid binding activity.

The lack of hydrophobic patches on the protein surface and the stability of the globular structure suggest several other possible explanations for lipid binding. For example, lipid binding may

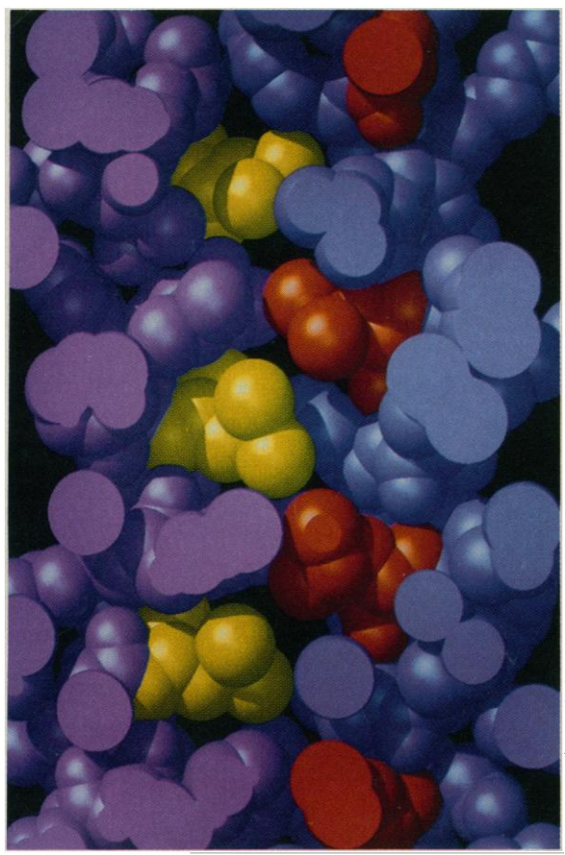
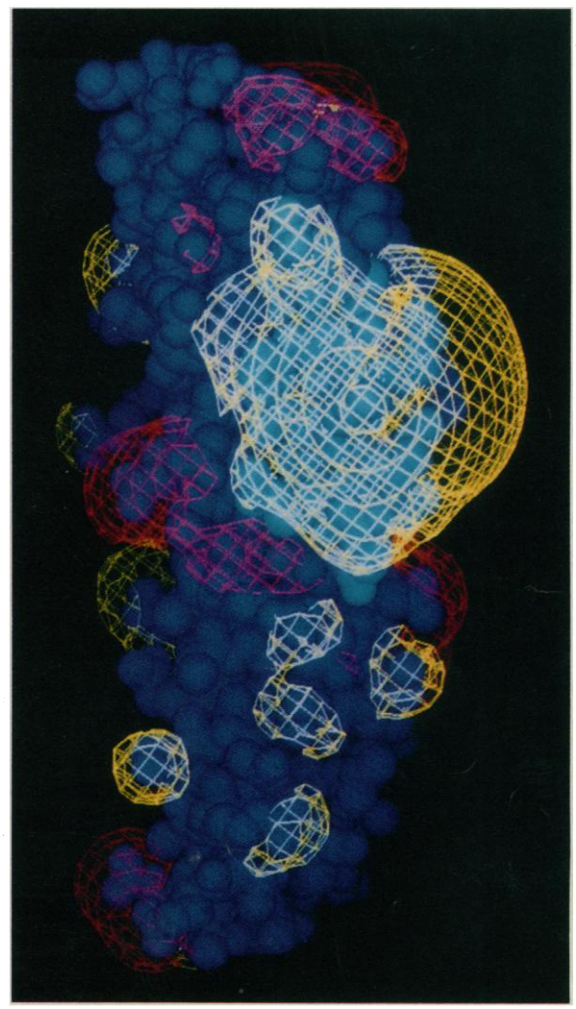


Fig. 5 (left). Cross section through a space-filling representation of the NH_2-terminal domain. Leucines on helix 1 (residues 30, 37, 43; yellow) and helix 4 (residues 133, 141, 148, 155; red) are arranged to form a "zipper" in a manner similar to that predicted by Landschulz *et al.* (*21*) for the C/EBP-type transcription factors. **Fig. 6 (right).** Electrostatic potential map of apoE. The DELPHI program (Biosym, San Diego, California) was used to calculate an approximate solution to the linearized Poisson-Boltzmann equation. The dielectric of the protein region (not including bound water molecules) was set to 2, while the solvent dielectric was set to 80. An ionic strength of 150 mM was assumed and only formal protein charges were included in the calculation. Positive (yellow) and negative (red) contours in the potential are evaluated at +2 and $-2\ kT/e^-$, respectively, and are shown together with a space-filling model of the entire NH_2-terminal domain (blue). Residues in the receptor-binding region (136–150) are colored cyan.



involve a major conformational change whereby the hydrophobic face of one or more helices would become accessible for interaction with nonpolar phospholipid tails. Alternatively, lipid binding could result from the interaction of charged surface amino acids and phospholipid head groups. Finally, it is possible that the missing disordered residues (1 to 22 and 167 to 191) become structured in the presence of lipid and mediate lipid binding. Previous biophysical studies have indicated that the 22-kD fragment of apoE does not undergo a detectable conformational change on lipid binding (*2*); a result that does not support the first hypothesis. Furthermore, we have found that the linker between the NH_2- and COOH-terminal domains becomes less protease-sensitive upon lipid binding (*37*). Although these data support a model in which lipids interact with the COOH-terminal region of the 22-kD domain and order it, further crystallographic and spectroscopic studies are required to elucidate the mechanism of lipid binding and its role in activating the binding of apoE to the LDL receptor.

Genetic variability in apoE. Three major isoforms of apoE have been characterized; these are apoE2, E3, and E4. Functional differences between these proteins appear to lead to quantitative changes in both plasma cholesterol and in the likelihood of coronary heart disease (*38*). Apolipoprotein-E3, the protein used for our study, is the most common isoform. The replacement of arginine by cysteine at position 158 (corresponding to the most common apoE2 isoform) results in defective LDL receptor binding (*39*) and is strongly linked to type III hyperlipoproteinemia, a genetic disorder associated with premature atherosclerosis. The mutation of Cys^{112} to Arg (apoE4) is associated with increased plasma cholesterol and LDL despite normal LDL receptor binding (*39*). In the absence of refined structures for these mutant proteins, we can attempt to use the structure of apoE3 to provide an understanding of the functional defects.

Arginine-158, site of the E2 mutation, is located near the COOH-terminal end of helix 4, well removed (>15 Å) from the other residues implicated in LDL receptor binding. The guanidinium group of Arg^{158} does not contribute directly to the large positive electrostatic potential surrounding the receptor-binding helix. Instead, it forms salt bridges with the acidic side chains of Glu^{96} and Asp^{154}, and as such may help to stabilize the pairing of helices 3 and 4. Previous experiments have suggested that the E2 mutation may induce a significant conformational change that reduces receptor binding (*40*). The finding that residue 158 lies far from the basic region of residues 136 to 150 yet has a dramatic effect on receptor binding is consistent with this hypothesis.

The Cys^{112} to Arg mutant (the apoE4 isoform) has full LDL receptor binding (*39*) but altered lipoprotein binding (strongly favoring VLDL over HDL) (*41*). This effect is surprising since the mutation occurs in the NH_2-terminal domain, yet lipoprotein binding is largely mediated by the COOH-terminal domain. The Cys^{112} is partially buried between helices 2 and 3, well isolated from the receptor-binding helix (Fig. 2C). Model building based on the apoE3 structure suggests that an arginine at this position could be easily accommodated by filling the solvent region surrounding the helix pair. In doing so, the arginine side chain may disrupt a specific interaction between the NH_2-terminal domain and the COOH-terminal domain, altering the structure or accessibility of the lipid binding determinants directly. Further crystallographic studies of the E2 and E4 isoforms should make it possible to test these hypotheses.

Note added in proof. Crystallographic studies of the *Locust migratoria* protein apolipophorin III have been reported since the submission of our manuscript (*42*). The structure of this protein is similar to that for apoE, suggesting that the elongated helical bundle may be a common structural motif for apolipoproteins. We have recently extended the resolution of the apoE data to 2.25 Å; coordinates for the newly refined structure will be deposited with the Brookhaven Protein Data Bank.

REFERENCES AND NOTES

1. R. W. Mahley, *Science* **240**, 622 (1988).
2. J. R. Wetterau, L. P. Aggerbeck, S. C. Rall, Jr., K. H. Weisgraber, *J. Biol. Chem.* **263**, 6240 (1988); L. P. Aggerbeck, J. R. Wetterau, K. H. Weisgraber, C.-S. C. Wu, F. T. Lindgren, *J. Biol. Chem.* **263**, 6249 (1988).
3. T. L. Innerarity, E. J. Friedlander, S. C. Rall, Jr., K. H. Weisgraber, R. W. Mahley, *ibid.* **258**, 12341 (1983).
4. S. H. Gianturco *et al.*, *ibid.*, p. 4526.
5. K. H. Weisgraber, *J. Lipid Res.* **31**, 1503 (1990).
6. M. S. Boguski, M. Freeman, N. A. Elshourbagy, J. M. Taylor, J. I. Gordon, *ibid.* **27**, 1011 (1986), M. S. Boguski, E. H. Birkenmeirer, N. A. Elshourbagy, J. M. Taylor, J. I. Gordon, *J. Biol. Chem.* **261**, 6398 (1986).
7. C.-C. Luo, W.-H. Li, M. N. Moore, L. Chan, *J. Mol. Biol.* **187**, 325 (1986).
8. H. J. Pownall, Q. Pao, D. Hickson, J. T. Sparrow, A. M. Gotto, Jr., *Biophys. J.* **37**, 175 (1982).
9. L. P. Aggerbeck, J. R. Wetterau, K. H. Weisgraber, R. W. Mahley, D. A. Agard, *J. Mol. Biol.* **202**, 179 (1988).
10. H. Hope, *Annu. Rev. Biophys. Biophys. Chem.* **19**, 197 (1990).
11. R. A. Crowther, in *Molecular Replacement Methods*, M. G. Rossmann, Ed. (Gordon & Breach, New York, 1971).
12. Refinement done with the CORELS program [J. L. Sussman, *Methods Enzymol.* **115**, 271 (1985)].
13. Refined with the HEAVY program [T. C. Terwilliger and D. Eisenberg, *Acta Cryst.* **A39**, 813 (1983)], modified by V. Ramalingam, UCSF.
14. B. C. Wang, *Methods Enzymol.* **115**, 90 (1985).
15. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987).
16. J. W. Ponder and F. M. Richards, *J. Mol. Biol.* **193**, 775 (1987).
17. S. R. Presnell and F. E. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6592 (1989).
18. S. Sheriff, W. A. Hendrickson, J. L. Smith, *J. Mol. Biol.* **197**, 273 (1987).
19. H. K. Das, J. McPherson, G. A. P. Bruns, S. K. Karathanasis, J. L. Breslow, *J. Biol. Chem.* **260**, 6240 (1985).
20. M. R. Kanost *et al.*, *ibid.* **263**, 10568 (1988).
21. N. Takahashi, Y. Takahashi, F. W. Putnam, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1906 (1985).
22. G. Lee, N. Cowan, M. Kirschner, *Science* **239**, 285 (1988).
23. S. A. Lewis, D. Wang, N. J. Cowan, *ibid.* **242**, 936 (1988).
24. J. P. Segrest, R. L. Jackson, J. D. Morrisett, A. M. Gotto, Jr., *FEBS Lett.* **38**, 247 (1974).
25. W. H. Landschulz, P. F. Johnson, S. L. McKnight, *Science* **240**, 1759 (1988).
26. E. K. O'Shea, R. Rutkowski, P. S. Kim, *ibid.* **243**, 538 (1989).
27. M. K. Gilson and B. H. Honig, *Nature* **330**, 84 (1987).
28. U. Beisiegel, W. Weber, G. Ihrke, J. Herz, K. K. Stanley, *ibid.* **341**, 162 (1989).
29. K. H. Weisgraber *et al.*, *J. Biol. Chem.* **258**, 12348 (1983).
30. W. A. Mann, R. E. Gregg, D. L. Sprecher, H. B. Brewer, Jr., *Biochim Biophys. Acta* **1005**, 239 (1989).
31. A. Lalazar *et al.*, *J. Biol. Chem.* **263**, 3542 (1988).
32. R. E. Pitas, T. L. Innerarity, R. W. Mahley, *ibid.* **255**, 5454 (1980).
33. T. J. Knott *et al.*, *Nature* **323**, 734 (1986); C.-Y. Yang *et al.*, *ibid.*, p. 738.
34. T. L. Innerarity, R. E. Pitas, R. W. Mahley, *J. Biol. Chem.* **254**, 4186 (1979).
35. J. R. Wetterau, L. P. Aggerbeck, K. H. Weisgraber, unpublished observations.
36. J. T. Sparrow, D. A. Sparrow, A. R. Culwell, A. M. Gotto, Jr., *Biochemistry* **24**, 6984 (1985).
37. K. H. Weisgraber, unpublished observations.
38. J. Davignon, R. E. Gregg, C. F. Sing, *Arteriosclerosis* **8**, 1 (1988).
39. K. H. Weisgraber, T. L. Innerarity, R. W. Mahley, *J. Biol. Chem.* **257**, 2518 (1982).
40. T. L. Innerarity, K. H. Weisgraber, K. S. Arnold, S. C. Rall, Jr., R. W. Mahley, *ibid.* **259**, 7261 (1984).
41. R. E. Gregg *et al.*, *J. Clin. Invest.* **78**, 815 (1986).
42. D. R. Breiter *et al.*, *Biochemistry* **30**, 603 (1991).
43. C. E. Kundrot and F. M. Richards, *Prot.: Struct. Function, Genet.* **3**, 71 (1988).

44. We thank J. Newdoll and the UCSF Computer Graphics Lab for help in preparing the figures. Supported by the Howard Hughes Medical Institute (D.A.A. and C.W.), by NIH grant HL-41633 (M.R.W., K.H.W. and R.W.M.); and by the Fannie and John Hertz Foundation (C.W.).

1 February 1991; accepted 17 May 1991