A mechanism for lipid binding to apoE and the role of intrinsically disordered regions coupled to domain–domain interactions

Carl Frieden[a,1], Hanliu Wang[b,2], and Chris M. W. Ho[c]

[a]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110; [b]Department of Chemistry, Washington University in St. Louis, St. Louis, MO 63130; and [c]Drug Design Methodologies LLC, St. Louis, MO 63103

Contributed by Carl Frieden, May 5, 2017 (sent for review March 28, 2017; reviewed by George D. Rose and Cheryl L. Wellington)

Relative to the apolipoprotein E (apoE) E3 allele of the *APOE* gene, apoE4 strongly increases the risk for the development of late-onset Alzheimer's disease. However, apoE4 differs from apoE3 by only a single amino acid at position 112, which is arginine in apoE4 and cysteine in apoE3. It remains unclear why apoE3 and apoE4 are functionally different. Described here is a proposal for understanding the functional differences between these two isoforms with respect to lipid binding. A mechanism is proposed that is based on the full-length monomeric structure of the protein, on hydrogen–deuterium exchange mass spectrometry data, and on the role of intrinsically disordered regions to control protein motions. It is proposed that lipid binds between the N-terminal and C-terminal domains and that separation of the two domains, along with the presence of intrinsically disordered regions, controls this process. The mechanism explains why apoE3 differs from apoE4 with respect to different lipid-binding specificities, why lipid increases the binding of apoE to its receptor, and why specific residues are conserved.

hydrogen–deuterium exchange | domain–domain interaction | conserved residues | protein structure | apolipoprotein E

ApoE is a 34-kDa protein whose normal function in the brain is to transport lipids and cholesterol to neuronal cells. In humans, there are three common isoforms—apoE2, apoE3, and apoE4—that appear to differ in important functional properties such as the distinct differences between apoE isoforms with respect to lipid transport (1). ApoE4 is the most studied since, relative to apoE3, it is known to be the major risk factor for the development of late-onset Alzheimer's disease (2, 3), whereas apoE2, the less common form, is associated with type III hyperlipoproteinemia (4, 5). A great many papers have discussed possible mechanisms for lipid binding (i.e., refs. 6–11), but there is no proposal for the mechanism of lipid binding that would explain the differences between the isoforms. From studies using FRET, small-angle X-ray diffraction, and electron paramagnetic resonance spectroscopy, it has been shown that apoE undergoes a conformational change on lipid binding, with the protein adopting a hairpin-like structure (7, 11–13). Here again, however, specific details are lacking.

In this paper, we assemble a number of experimental and computational observations to provide a model for the mechanism of lipid binding. This model includes a consideration of the full-length apoE structure, results from hydrogen–deuterium exchange (HDX) experiments, information on conserved and nonconserved residues, an appreciation of the role of intrinsically disordered regions (IDRs), and molecular dynamics calculations.

The proposed model explains why there are differences between apoE isoforms with respect to lipid binding and why lipid enhances apoE binding to receptors such as the low-density lipoprotein receptor (LDLR). As such, it opens the way to develop small molecular weight compounds that could preferentially influence the behavior of the apoE isoforms.

Results and Discussion

The ApoE Structure. The molecular mass of monomeric apoE is 34 kDa, but determination of the full-length structure of the protein has been difficult. At concentrations below 10 μM, the protein forms tetramers in the absence of lipids (14), whereas at higher concentrations the protein forms aggregates (15), thus precluding both crystallographic and NMR measurements.

In general, the protein is described as having an N- and C-terminal domain, with the C-terminal domain being involved in tetramer formation. Because the N-terminal domain does not form higher molecular weight oligomers, its structure was determined some years ago (16). More recent studies have used both NMR and X-ray measurements [Protein Data Bank (PDB) ID codes 2KCE, 1GS9, 1B68, and 1NFN]. From these experiments enough data were obtained to show that there is a hinge region between the N- and C-terminal domains.

In 2004, Fan et al. (17) found that making five mutations in the isolated C-terminal peptide resulted in a monomeric structure. Using these same five mutations in full-length apoE3 yielded a monomeric form that was biologically active (18). In 2011, Chen et al. (10), again using the same five mutations, were able to determine an NMR structure of apoE3 by segmentally labeling the N-terminal domain of the protein with ^{15}N and ^{13}C and the C-terminal domain with ^{15}N. Consequently, they were able to determine an NMR structure of the full-length protein, the first, and currently the only, such structure.

With respect to the interaction between the N- and C-terminal domain, the structure determined by Chen et al. (10) was at odds

Significance

Alzheimer's disease is the most common cause of dementia. It is irreversible, devastating, and costly. Apolipoprotein E4 (ApoE4), a protein whose function in the brain is to transport cholesterol-laden lipid, is a major risk factor for the development of late-onset Alzheimer's disease relative to the common apoE3 isoform. The two proteins differ by only a single amino acid. There has been no detailed mechanism explaining how apoE functions. Here, we propose a mechanism for lipid binding that involves simple domain–domain movement coupled to the role of intrinsically disordered regions of the protein. This mechanism may lead to the development of small molecular weight compounds that interfere with the role apoE4 plays in the development of Alzheimer's disease.

Author contributions: C.F. designed research; H.W. and C.M.W.H. performed research; and C.F. wrote the paper.

Reviewers: G.D.R., Johns Hopkins University; and C.L.W., University of British Columbia.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. Email: friedenc@wustl.edu.

[2]Present address: Analytical Research and Development, Pfizer Inc., Chesterfield, MO 63017.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1705080114/-/DCSupplemental.

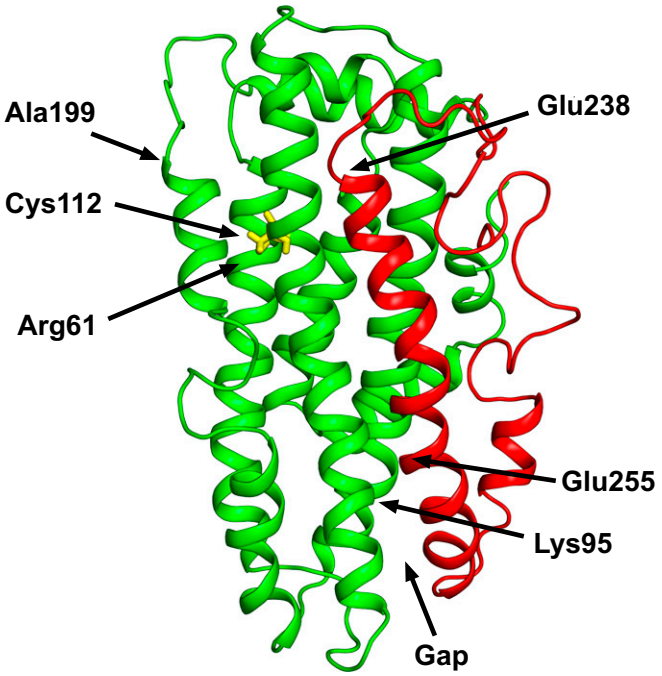

Fig. 1. An average NMR structure of full-length monomeric apoE3 as determined by Chen et al. (10) (PDB ID code 2L7B). In this depiction, the five residues that were mutated to prevent oligomer formation have been back mutated to their original amino acids using PyMol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC). The C-terminal domain is colored red. The gap between the N and C termini is indicated. The positions of Cys112, Lys95, Arg61, and Glu255 are shown, as are the end of the N-terminal domain (Ala199) and the beginning of the C-terminal domain (Glu238). As noted in the text, the authors proposed a salt bridge involving Glu255 to Lys95 rather than Arg61, as had been suggested earlier. This and other figures showing the apoE3 structure were generated using PyMol.

with an earlier prediction that had proposed only a single salt bridge between the two domains of apoE4 and no such interactions in apoE3 (9). In addition, Chen et al. (10) found that the C-terminal domain was not completely helical, as had been proposed, but contained a large IDR. The earlier predicted model had suggested the importance of a hinge region between the N and C termini but failed to recognize the multiple interactions between the domains. Some recent papers continue to use this model to relate structure to function (19) rather than using the full-length NMR structure determined by Chen et al. (10). In some ways, as discussed here, the earlier model was correct in that domain movement is in fact critical for the function of apoE. However, the details of the predicted structure were incorrect, as was the suggested salt bridge, precluding an understanding of the function of apoE and, in particular, an understanding of lipid binding.

Fig. 1 shows an average NMR structure of monomeric apoE3 determined by Chen et al. (10) (PDB ID code 2L7B). As indicated in the figure, we define the N-terminal domain as residues 1–198, the C-terminal domain as residues 238–299, and the hinge helix as residues 199–237 (10). What is absolutely critical to the proposal made below is the interaction between the N- and C-terminal domains. Base on mutational data, Dong and Weisgraber (20) had proposed a single salt bridge in apoE4 between Arg61 (in helix 2 of the N-terminal domain) and Glu255 (in the C-terminal domain). In the full-length NMR structure of mutated apoE3 determined by Chen et al., (10) however, there are many interactions between the N and C termini. Specifically, these authors indicated that Glu255 forms a salt bridge with Lys95 located in helix 3 of the N-terminal domain rather than the predicted salt bridge between Glu255 and Arg61 of helix 2 in apoE4 (10). Although this difference may appear to be trivial, it completely changes the relationship between the N- and C-terminal domains and is crucial for understanding lipid binding. Furthermore, based on the data presented below, it seems likely that essentially the same multiple interactions between the N- and C-terminal domains occur for all isoforms of apoE.

Structural Differences Between ApoE3 and ApoE4. HDX coupled to mass spectrometry (HDX-MS) is now a common method used to examine local hydrogen bond networking and solvent accessibility of backbone amide protons (21). Fig. 2 shows HDX data comparing wild-type apoE3 and apoE4. In this figure, only peptides that show differences between apoE3 and apoE4 are presented. All of the peptides from apoE4 exchange more rapidly than those in apoE3, suggesting that in these specific regions apoE4 is more dynamic than apoE3. The remaining peptides, covering the full-length sequence, show no HDX differences. The complete data are shown in Fig. S1 *A–D*. Heatmaps of wild-type apoE4 and apoE3 are shown in Fig. S1 *E* and *F* for data visualization. The location of the α-helices is consistent with a recent extensive HDX study of Chetty et al. (22) that focused on the dynamics and stability of the helical segments of apoE3. The experiments carried out in the paper of Chetty et al. (22) did compare wild-type apoE3 and apoE4, but there were some differences. Their comparison experiments were performed at 5 °C, whereas ours were at 25 °C. They also relied on kinetic fitting results for comparison, whereas we used raw HDX data. Both our HDX study and those of Chetty et al. (22) support the validity of the NMR structure of the apoE3 monomeric mutant as determined by Chen et al. (10). Any differences in comparing wild-type apoE3 and apoE4 between our data and those of Chetty et al. (22) are likely due to different sample handling and data analysis methods.

As shown by the data in Fig. 2, the largest differences between apoE3 and apoE4 appear to be in peptides 27–37, 60–64, and 102–108. Arg61, within the peptides 60–64, has long been known to affect lipoprotein preferences (23) and, as mentioned above, was believed to form a salt bridge with Glu255 in apoE4. Instead, it is close to Cys112 (Fig. 1). All of the peptides that differ in HDX behavior (except for peptides 148–159) are actually quite close in space to position 112, the site of the cysteine/arginine change between apoE4 (Arg112) and apoE3 (Cys112). Although these experiments reveal peptides that differ in exchange properties, they do not specify which are the critical residue(s). Nor is it clear from these data what the extent of the structural change is. Despite that, these structural changes may define why apoE3 and apoE4 have different lipid specificities as discussed later.

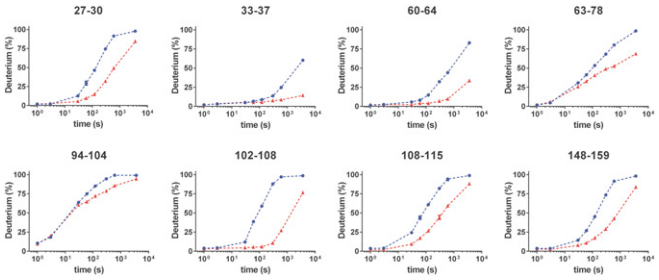

Fig. 2. Structural differences between wild-type apoE3 and apoE4 as determined by HDX-MS studies plotted as a function of log time (s). The blue line represents data for apoE4, whereas the red line represents data for apoE3. Only those regions that show differences are shown here. The complete data are given in Fig. S1.

IDRs in Proteins: Domain Motions, Hinge Regions, and Smaller Segmental Motions. IDRs allow structured portions of a protein to move relative to one another. The different extent of motions of domains and of individual helices, however, is unquestionably the basis for understanding protein function, as has been recently reviewed (24–26).

ApoE3, as determined by Chen et al. (10), contains both IDRs and smaller flexible regions that surround structured helices. Approximately a third of the protein can be defined this way. Many investigators have noted that the helix defined by residues 210–223 has large IDRs surrounding it (residues 199–209 and 224–239). Although this critical hinge region controls the interaction between the N- and C-terminal domains, what has not been fully appreciated is that IDRs and flexible regions also occur within the N-terminal domain itself. Fig. 3 shows the NMR apoE3 structure with the IDRs and flexible regions emphasized (in red). What is apparent from this figure is that many helices, both short and long (colored blue), are surrounded by regions that are similar to that which controls the relative motions of the N and C termini.

With IDRs and flexible regions, it is almost certain that apoE senses many environments as domains and segments move. The mechanism of lipid binding as proposed here requires that these motions exist and that the many different conformational forms occur with different equilibrium constants.

Equally notable is that the protein has numerous salt bridges. Charged residues (Arg, Lys, Glu, Asp) account for one-third of the sequence. Their role in affecting stability or regions of flexibility is currently unclear, although their widespread surface distribution may steer lipids away from incorrect binding sites.

Conserved Regions and Residues. ApoE is the product of the *APOE* gene. By definition, and irrespective of other sequence differences, apoE4 has arginines at positions 158 and 112, whereas apoE3 has a cysteine at position 112 and apoE2 has cysteines at both 112 and 158. By this definition, all mammalian species, except humans, express only apoE4. Thus, one can determine highly conserved residues from data already available. A previous paper discussed those amino acid residues that were conserved among ~60 different mammalian species (27). This study was conducted as a way to find specific ligand-binding sites as well as sites related to apoE4 function. Of the 299 amino acids in apoE, only ~15% are absolutely conserved, another 12% have a single amino acid change (among the ~60 species), and 7% have two changes (27). Even assuming some errors in determining protein sequences from DNA databases, at least 65% of residues are variable, with the remaining 35% being highly conserved. The data for all sequences of apoE4 are given in the Dataset S1.

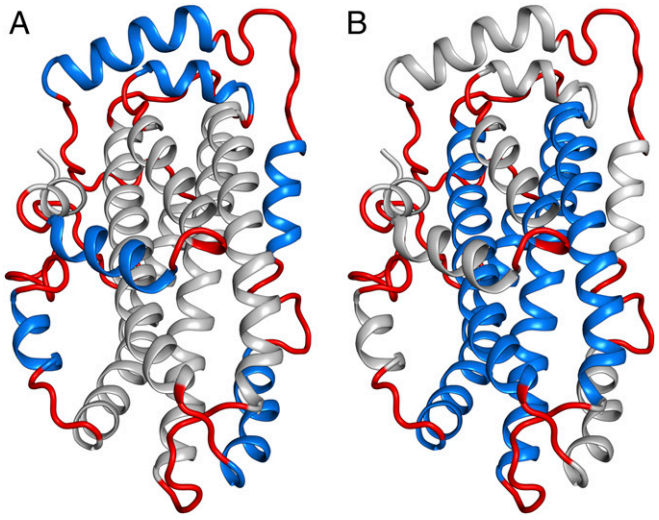

Fig. 3. A representation of the full-length structure of apoE3 with the IDRs and flexible regions (in red) surrounding helical regions. (*A*) Short helices shown in blue. (*B*) Long helices shown in blue. The hinge connecting the N- and C-terminal domains is at the top of the figure. These regions comprise ~30% of the structure.

Table 1. Conserved residues* around the putative lipid-binding site

Helix 3, 92–104	Helix 4, 150–165	C-terminal, 253–267
Arg92	Arg150	Gln253
Lys95	Asp151	Ala254
Glu96	Asp154	Phe257
Gln98	Arg158	Arg260
Ala100	Leu159	Trp264
Gln101	Ala160	Phe265
Arg103	Tyr162	Glu266
Leu104	Gly165	Pro267

*Residues with only one difference in 63 species are considered conserved. The complete list of conserved and nonconserved residues is given in Dataset S1.

Among strongly conserved regions are many residues from 253 to 267 in the C-terminal domain as well as residues in regions 92–104 and 150–165 in the N-terminal domain. Table 1 lists conserved residues in these regions. Fig. 4 shows that those regions surround the gap shown in Fig. 1 and that many may be involved in salt bridges. Also of note is that the region 150–165 of the N-terminal domain includes Arg158, the residue that is changed to cysteine in apoE2.

Because the function of apoE is to transport lipid and cholesterol, one would expect that the residues involved would be highly conserved. As discussed below, we postulate that these three regions listed in Table 1, from both the C- and N-terminal domains and specifically those around Trp264 and Ser94 (28), are involved with lipid binding.

Based on the structure of apoE3 determined by Chen et al. (10), the region from residues 278–299 is intrinsically disordered. As pointed out in an earlier publication (27), this region shows little conservation of residues among different mammals. One can speculate that, rather than random drift, these differences relate to dietary lipid intake or differences in lipid metabolism, especially in different mammalian species.

Lipid Binding. Defining lipid binding has not been trivial, and many mechanisms have been proposed (5, 7, 10, 28–32). Essentially all involve the C-terminal domain, which is also involved in the oligomerization of wild-type apoE. Consequently, only monomeric apoE binds lipid (33). Even ignoring this complication, it is unlikely that binding is a simple one-step process. It has been suggested, for example, that there is an initial binding to the C-terminal domain followed by a second step involving binding to the N-terminal domain (28, 31, 34). Indeed, Nguyen et al. suggest that the overall stability of domains in the entire molecule influences lipid binding (35), meaning that lipid binding involves residues from both the N- and C-terminal domains, residues that are close in the apoE structure but sequentially distant. Various proposals also reflect the fact that apoE must be able to accommodate many different lipoprotein particles with different compositions and sizes and still show some specificity between different isoforms (5). IDRs may account for this ability. It should be noted in this regard that many publications use the interaction of apoE with plasma lipids even though lipid composition in the brain is quite different (36).

Steps in Lipid Binding. In the model proposed here, the initial lipid-binding step does not result in a reorganization of any

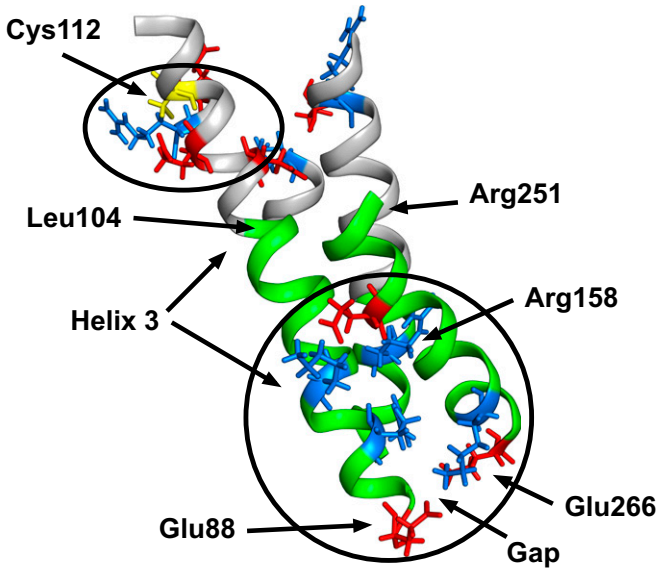

Fig. 4. Regions of apoE3, shown in green, that surround the gap between the N- and C-terminal domains. Shown within this gap are residues involved in salt bridges as proposed by Chen et al. (10) colored either blue (positively charged) or red (negatively charged). Arg158, which is a cysteine in apoE2, is labeled. Regions that may encompass the lipid-binding site are from Glu88 to Leu104 in the N-terminal domain and from Arg251 to Glu266 in the C-terminal domain. Also shown are charged residues around position 112 that, based on the data shown in Fig. 2, may be affected by the arginine residue at position 112 in apoE4. As discussed in the text, a portion of Helix 3 links the N-terminal portion of the lipid-binding site to the region of the cysteine to arginine difference at position 112. Residues that could be affected by this change, as indicated by Fig. 2, are shown.

portion of the molecule but rather exploits simple domain movements controlled by the hinge region between residues Leu199 and Glu238, as shown in Fig. 1.

Between the N- and C-terminal domains, there is a gap that we propose is the site of the initial step for lipid binding (shown in Figs. 1 and 4). As indicated in Fig. 4, the separation of the N- and C-terminal domains in this region appears to be controlled by several salt bridges: an Arg/Glu salt bridge at the mouth of the opening and at least two buried more deeply. Residues pointing into this gap are both charged (Glu88, Arg92, Lys95, Glu96, Arg103, Glu255, and Glu266) and uncharged (Thr89, Ala99, Ala259, and Ser263). Fig. 4 shows those regions of apoE that surround the lipid-binding site (Glu88 to Leu104 and Arg251 to Glu266) as well as charged residues pointing into the gap. Arg158, which in apoE2 is a cysteine, is directly behind the lipid-binding site and is surrounded by flexible regions. Fig. 4 also shows that a portion of Helix 3 links the lipid-binding site to position 112.

Thus, what may be a first step in lipid binding is domain–domain separation between the N- and C-terminal domains where salt bridges are broken by the charged lipid molecule. This simple domain–domain opening allowing lipid binding has not been described previously and rests upon the structure proposed by Chen et al. (10). A video of this separation of domains, based on using molecular dynamics, is presented in Movie S1. As discussed below, this proposal is also consistent with the structural differences between apoE3 and E4 and, importantly, with the observation that lipidated apoE binds more tightly to a receptor than does nonlipidated apoE. How far the C and N termini move apart remains unclear at this time, but the movement of the domains exposes the receptor-binding site as discussed below.

As mentioned above, the arginine at position 158 in apoE3 and apoE4 is changed to a cysteine in apoE2. Based on the structure of apoE3, the arginine side chain projects into the gap between the N- and C-terminal domains. The fact that this residue is at the mouth of the putative lipid-binding site may affect apoE2 binding to lipid and subsequently to the LDL receptor (37–39), as the gap between the N and C termini may collapse with the smaller cysteine side chain at position 158.

As noted above, apoE must be able to accommodate lipoprotein particles of different compositions and sizes. Clearly a lipoprotein particle may be much larger than any space available within the apoE molecule. Therefore, apoE itself must influence the behavior of the lipoprotein particle. There may be several ways to do this. For example, the protein may affect the curvature of the lipoprotein particle, allowing a small region to intercalate into the space between the N- and C-terminal domains. There have been many investigations to study this particular process as reviewed by Stachowiak et al. (40, 41) and McMahon and Gallop (42). Although this process is a crucial part of the proposed mechanism for apoE, it is unclear what it might involve.

For larger particles, the separation of the N and C domains may be so large that helices may lie flat on the lipid surface, as suggested by Schneeweis et al. (29). The large number of flexible regions surrounding helices, as shown in Fig. 3, may allow an extensive movement to expose different residues for different lipids.

The Receptor-Binding Site. Many investigations have shown that the LDLR site on apoE is located within the region of residues 140–150 (i.e., ref. 37). Lalazar et al. (43) specifically noted Lys143 as important, whereas Lund-Katz et al. (44) pointed out the involvement of both Lys143 and Lys146. This region is highly conserved (27) and in the wild-type apoE3 structure is solvent-inaccessible based on the GETAREA program (45). It is well-known that the lipidated form of apoE binds to the LDLR much more readily than does the nonlipidated protein. As noted by Chen et al. (10), the major LDLR-binding region is shielded by the C-terminal domain. Molecular dynamics can be used to separate the N- and C-terminal domains and thus emulate lipid binding as proposed in the model. Table 2 shows that on moving the N and C termini apart, residues 142, 143 146, and 147 become solvent-accessible, with the largest changes between the closed and open forms being residues 143 and 146, in agreement with the experimental data. A video of the domain–domain separation is presented in Movie S1.

The location of the receptor site in the model proposed here is directly above those N- and C-terminal domains predicted to separate on lipid binding. This easily explains why lipidated apoE binds lipid more effectively to receptor relative to the nonlipidated protein.

Why Do ApoE3 and ApoE4 Differ in Lipoprotein Particle Specificity? It has been appreciated for many years that apoE isoforms exhibit different specificities for different plasma lipoprotein particles.

Table 2. Solvent accessibility of residues 140–148 for apoE3 before and after domain–domain separation

Residue	Closed	Open*
His140	2.7	8.9
Leu141	3.3	5.0
Arg142	0.4	43.0
Lys143	8.8	66.9
Leu144	1.0	0.1
Arg145	7.3	5.2
Lys146	3.3	62.3
Arg147	0.4	42.2
Leu148	8.8	1.2

Data are based on the GetArea program (45). Numbers are the ratio of side-chain surface area to "random coil" value per residue. Those values larger than 50 are considered to be surface exposed, whereas those less than 20 are not.

*Data taken from the final frame of an MD calculation.

ApoE binds tightly to lipoprotein particles such as VLDL and HDL but with different specificities in that apoE4 binds preferentially to VLDL whereas apoE3 binds preferentially to HDL (9). Nguyen et al. (34, 46) suggest that different preferences of apoE3 binding to HDL arise because binding is mediated primarily by both the lipid particle surface and differences in organization of apoE. Although the specific residues that are structurally different as determined by HDX are not defined in Fig. 2, it is clear that there are structural differences as a consequence of whether cysteine or arginine is located at position 112. Furthermore, these differences are proximal to any lipid molecules that may be present. This difference may be the reason that apoE3 and apoE4 have different preferences for lipoprotein particles of different sizes.

Domain Rotation May also Affect Specificity in Lipid Binding. Also a possible reason for differences in lipoprotein particle binding is the great flexibility imparted to the protein as a consequence of IDRs. As shown by Fig. 3, the IDRs connecting the large hinge region to the N- and C-terminal domains are relatively long: 10 amino acids in the N-terminal domain and 15 amino acids in the C-terminal domain. In addition to being crucial with respect to opening the structure to allow lipoprotein particles to bind, the long IDRs might also allow rotation of domains relative to their position in the unopened protein. This could have several consequences. For example, the long unstructured region in the C-terminal domain (residues 278–299) is relatively hydrophobic, whereas the adjacent helix (residues 237–265) is highly charged. The 237–265 helix is amphipathic, and rotation of this domain would present a totally different environment to the lipoprotein particle. Although both regions, 237–265 and 278–299, are not well conserved, the disordered region from 278 to 299 at the end of the C-terminal domain shows great variability throughout the different mammalian species, perhaps acting as another mechanism for binding different lipoprotein particles.

Other Lipid-Binding Functions of ApoE. ApoE is widely distributed throughout the body. As pointed out in many reviews (1, 5, 9, 47–50), apoE has multiple lipid-binding properties. The central nervous system contains only HDL-like lipoproteins, and apoE is found on these particles there. In the plasma, however, apoE is associated with VLDL, chylomicron remnants, and some HDL particles. One can reasonably ask whether the mechanism proposed here is also applicable to binding such a large number of different lipoprotein particles between the brain and the plasma. The answer may lie in the observation of the large number of IDRs and flexible regions that are present in apoE (Fig. 3) that allow many different surfaces to be exposed. As discussed above, IDRs allow large regions within any given domain to rotate without extensive unfolding (and subsequent refolding) to some different conformation. In addition, regions within any given domain may reorient smaller regions of the protein. It seems quite possible that the mechanism proposed here for binding to lipids in the brain may also be applicable to lipid binding in the plasma.

Conclusion

A summary of the proposal made herein is shown in Fig. 5. What is satisfying about this proposal is how the properties of apoE with respect to lipid binding can be easily explained. First, the proposal assumes that the multiple interactions between the N- and C-terminal domains are essentially the same for all apoE isoforms. Thus, the mechanism for lipid binding involves simple changes in domain–domain interaction, a common mode of behavior for many proteins. Importantly, it is not necessary to assume a large rearrangement of individual protein segments. Second, structural differences between apoE3 and apoE4 are located in regions where one might expect that the cysteine-to-arginine change can define binding and specificity. In a previous paper, Frieden (27) noted the conservation of residues 34–41 and 59–63 and stated it was not clear why these were conserved. Now it is clear that they reflect structural differences between apoE3 and apoE4. Third, those regions that define the function of apoE become accessible on movement of the N- and C-terminal domains. Of particular interest is the receptor-binding region that is not solvent-accessible in nonlipidated apoE but becomes accessible to solvent on moving the two domains apart. This easily explains why lipidated apoE binds much better to receptors relative to the nonlipidated protein. Fourth, the mechanism provides an explanation for why so many IDRs exist: to allow the protein to carry out multiple functions with respect to lipid binding. Finally, it is important to note the role of conserved residues. Because the function of apoE is to bind and transport lipids, one expects those residues to be conserved. Of the residues that are not conserved, it seems reasonable to assume that different mammalian species handle lipids differently, especially in the plasma.

The proposal here opens many opportunities for experimental testing. If validated, it will allow development of therapeutic agents that could affect the behavior of apoE4 relative to apoE3. It could also lead to small molecular weight compounds that could block the binding of specific lipids.

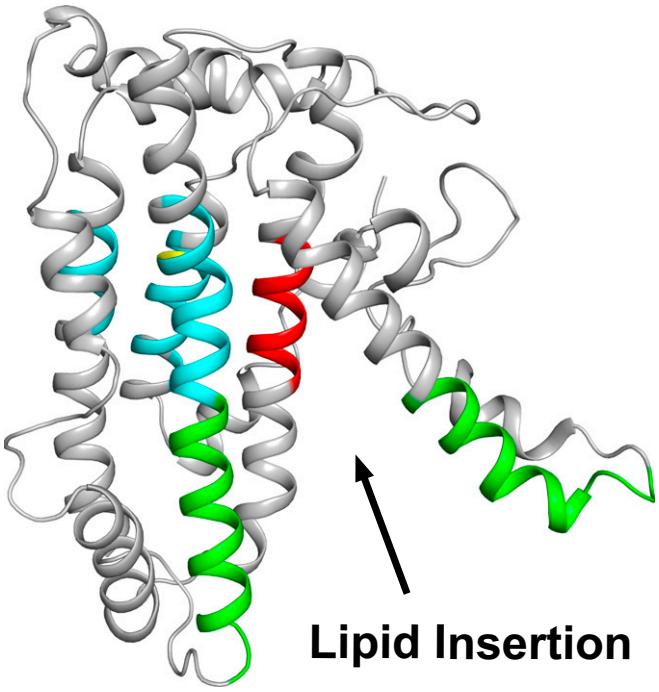

Fig. 5. Snapshot of domain–domain separation on lipid binding to apoE3. The structure was obtained using molecular dynamics to move apart the N- and C-terminal domains. Shown in green are portions of regions involved in lipid binding. In red is the apoE3 receptor-binding site and in blue are regions of structural differences based on HDX experiments. Cysteine112 (yellow) is within this region. Movie S1 provides a video moving from the closed to a completely open form.

Materials and Methods

ApoE isoforms were prepared as described by Mondal et al. (51). HDX-MS was performed as described in *SI Materials and Methods*. Continuous HDX was conducted similar to a previously described protocol (51). All molecular dynamics simulations were run at 300K using the AMBER ff03 force field (52) with the TIP3P explicit solvent model (53).

ACKNOWLEDGMENTS. The authors acknowledge important discussions with Drs. David Holtzman, Melissa Brereton, and Gregory DeKoster for help with figures. H.W. supported by NIH Grant P41GM103422. We thank Dr. Michael Gross for the use of HDX-MS equipment provided by NIH Grant 1S10OD016298 and Berevan Baban for preparations of apoE isoforms. This work is supported by NIH NIA Grant RF1AG044331 (to C.F.).

1. Hauser PS, Narayanaswami V, Ryan RO (2011) Apolipoprotein E: From lipid transport to neurobiology. *Prog Lipid Res* 50:62–74.
2. Corder EH, et al. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261:921–923.
3. Saunders AM, et al. (1993) Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 43:1467–1472.
4. Dong LM, et al. (1996) Novel mechanism for defective receptor binding of apolipoprotein E2 in type III hyperlipoproteinemia. *Nat Struct Biol* 3:718–722.
5. Phillips MC (2014) Apolipoprotein E isoforms and lipoprotein metabolism. *IUBMB Life* 66:616–623.
6. Hauser PS, Ryan RO (2013) Impact of apolipoprotein E on Alzheimer's disease. *Curr Alzheimer Res* 10:809–817.
7. Peters-Libeu CA, Newhouse Y, Hatters DM, Weisgraber KH (2006) Model of biologically active apolipoprotein E bound to dipalmitoylphosphatidylcholine. *J Biol Chem* 281:1073–1079.
8. Peters-Libeu CA, Newhouse Y, Hall SC, Witkowska HE, Weisgraber KH (2007) Apolipoprotein E*dipalmitoylphosphatidylcholine particles are ellipsoidal in solution. *J Lipid Res* 48:1035–1044.
9. Hatters DM, Peters-Libeu CA, Weisgraber KH (2006) Apolipoprotein E structure: Insights into function. *Trends Biochem Sci* 31:445–454.
10. Chen J, Li Q, Wang J (2011) Topology of human apolipoprotein E3 uniquely regulates its diverse biological functions. *Proc Natl Acad Sci USA* 108:14813–14818.
11. Hatters DM, Voss JC, Budamagunta MS, Newhouse YN, Weisgraber KH (2009) Insight on the molecular envelope of lipid-bound apolipoprotein E from electron paramagnetic resonance spectroscopy. *J Mol Biol* 386:261–271.
12. Hatters DM, Peters-Libeu CA, Weisgraber KH (2005) Engineering conformational destabilization into mouse apolipoprotein E. A model for a unique property of human apolipoprotein E4. *J Biol Chem* 280:26477–26482.
13. Drury J, Narayanaswami V (2005) Examination of lipid-bound conformation of apolipoprotein E4 by pyrene excimer fluorescence. *J Biol Chem* 280:14605–14610.
14. Garai K, Frieden C (2010) The association–dissociation behavior of the ApoE proteins: Kinetic and equilibrium studies. *Biochemistry* 49:9533–9541.
15. Perugini MA, Schuck P, Howlett GJ (2000) Self-association of human apolipoprotein E3 and E4 in the presence and absence of phospholipid. *J Biol Chem* 275:36758–36765.
16. Wilson C, Wardell MR, Weisgraber KH, Mahley RW, Agard DA (1991) Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E. *Science* 252:1817–1822.
17. Fan D, Li Q, Korando L, Jerome WG, Wang J (2004) A monomeric human apolipoprotein E carboxyl-terminal domain. *Biochemistry* 43:5055–5064.
18. Zhang Y, et al. (2007) A monomeric, biologically active, full-length human apolipoprotein E. *Biochemistry* 46:10722–10732.
19. Huang Y, Mahley RW (2014) Apolipoprotein E: Structure and function in lipid metabolism, neurobiology, and Alzheimer's diseases. *Neurobiol Dis* 72(Pt A):3–12.
20. Dong LM, Weisgraber KH (1996) Human apolipoprotein E4 domain interaction. Arginine 61 and glutamic acid 255 interact to direct the preference for very low density lipoproteins. *J Biol Chem* 271:19053–19057.
21. Weis DD (2016) *Hydrogen Exchange Mass Spectrometry of Proteins: Fundamentals, Methods, and Applications* (John Wiley & Sons, Inc., Chichester, UK).
22. Chetty PS, Mayne L, Lund-Katz S, Englander SW, Phillips MC (2017) Helical structure, stability, and dynamics in human apolipoprotein E3 and E4 by hydrogen exchange and mass spectrometry. *Proc Natl Acad Sci USA* 114:968–973.
23. Dong LM, et al. (1994) Human apolipoprotein E. Role of arginine 61 in mediating the lipoprotein preferences of the E3 and E4 isoforms. *J Biol Chem* 269:22358–22365.
24. Delaforge E, et al. (2016) Investigating the role of large-scale domain dynamics in protein-protein interactions. *Front Mol Biosci* 3:54.
25. Palamini M, Canciani A, Forneris F (2016) Identifying and visualizing macromolecular flexibility in structural biology. *Front Mol Biosci* 3:47.
26. Necci M, Piovesan D, Tosatto SCE (2016) Large-scale analysis of intrinsic disorder flavors and associated functions in the protein sequence universe. *Protein Sci* 25: 2164–2174.
27. Frieden C (2015) ApoE: The role of conserved residues in defining function. *Protein Sci* 24:138–144.
28. Mizuguchi C, et al. (2012) Fluorescence analysis of the lipid binding-induced conformational change of apolipoprotein E4. *Biochemistry* 51:5580–5588.
29. Schneeweis LA, Koppaka V, Lund-Katz S, Phillips MC, Axelsen PH (2005) Structural analysis of lipoprotein E particles. *Biochemistry* 44:12525–12534.
30. Lu B, Morrow JA, Weisgraber KH (2000) Conformational reorganization of the four-helix bundle of human apolipoprotein E in binding to phospholipid. *J Biol Chem* 275: 20775–20781.
31. Fisher CA, Ryan RO (1999) Lipid binding-induced conformational changes in the N-terminal domain of human apolipoprotein E. *J Lipid Res* 40:93–99.
32. Saito H, et al. (2003) Domain structure and lipid interaction in human apolipoproteins A-I and E, a general model. *J Biol Chem* 278:23227–23232.
33. Garai K, Baban B, Frieden C (2011) Dissociation of apolipoprotein E oligomers to monomer is required for high-affinity binding to phospholipid vesicles. *Biochemistry* 50:2550–2558.
34. Nguyen D, Dhanasekaran P, Phillips MC, Lund-Katz S (2009) Molecular mechanism of apolipoprotein E binding to lipoprotein particles. *Biochemistry* 48:3025–3032.
35. Nguyen D, et al. (2014) Influence of domain stability on the properties of human apolipoprotein E3 and E4 and mouse apolipoprotein E. *Biochemistry* 53:4025–4033.
36. Wellington CL, Frikke-Schmidt R (2016) Relation between plasma and brain lipids. *Curr Opin Lipidol* 27:225–232.
37. Weisgraber KH (1994) Apolipoprotein E: Structure-function relationships. *Adv Protein Chem* 45:249–302.
38. Yamamoto T, Choi HW, Ryan RO (2008) Apolipoprotein E isoform-specific binding to the low-density lipoprotein receptor. *Anal Biochem* 372:222–226.
39. Mahley RW, Weisgraber KH, Huang Y (2009) Apolipoprotein E: Structure determines function, from atherosclerosis to Alzheimer's disease to AIDS. *J Lipid Res* 50: S183–S188.
40. Stachowiak JC, Hayden CC, Sasaki DY (2010) Steric confinement of proteins on lipid membranes can drive curvature and tubulation. *Proc Natl Acad Sci USA* 107: 7781–7786.
41. Stachowiak JC, Brodsky FM, Miller EA (2013) A cost-benefit analysis of the physical mechanisms of membrane curvature. *Nat Cell Biol* 15:1019–1027.
42. McMahon HT, Gallop JL (2005) Membrane curvature and mechanisms of dynamic cell membrane remodelling. *Nature* 438:590–596.
43. Lalazar A, et al. (1988) Site-specific mutagenesis of human apolipoprotein E. Receptor binding activity of variants with single amino acid substitutions. *J Biol Chem* 263: 3542–3545.
44. Lund-Katz S, et al. (2001) Effects of polymorphism on the microenvironment of the LDL receptor-binding region of human apoE. *J Lipid Res* 42:894–901.
45. Fraczkiewicz R, Braun W (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comput Chem* 19: 319–333.
46. Nguyen D, et al. (2010) Molecular basis for the differences in lipid and lipoprotein binding properties of human apolipoproteins E3 and E4. *Biochemistry* 49: 10881–10889.
47. Mahley RW, Rall SC, Jr (2000) Apolipoprotein E: Far more than a lipid transport protein. *Annu Rev Genomics Hum Genet* 1:507–537.
48. Getz GS, Reardon CA (2009) Apoprotein E as a lipid transport and signaling protein in the blood, liver, and artery wall. *J Lipid Res* 50:S156–S161.
49. Liu CC, Liu CC, Kanekiyo T, Xu H, Bu G (2013) Apolipoprotein E and Alzheimer disease: Risk, mechanisms and therapy. *Nat Rev Neurol* 9:106–118.
50. Giau VV, Bagyinszky E, An SS, Kim SY (2015) Role of apolipoprotein E in neurodegenerative diseases. *Neuropsychiatr Dis Treat* 11:1723–1737.
51. Mondal T, et al. (2016) ApoE: In vitro studies of a small molecule effector. *Biochemistry* 55:2613–2621.
52. Kollman P (1996) Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules. *Acc Chem Res* 29:461–469.
53. Jorgensen WL, Madura JD (1983) Quantum and statistical studies of liquids. 25. Solvation and conformation of methanol in water. *J Am Chem Soc* 105:1407–1413.
54. Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* 17:75–86.
55. Essmann U, et al. (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103: 8577–8593.
56. Hess B (2008) P-LINCS: A parallel linear constraint solver for molecular simulation. *J Chem Theory Comput* 4:116–122.
57. Feenstra KA, Hess B, Berendsen HJC (1999) Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J Comput Chem* 20: 786–798.
58. Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. *J Chem Phys* 126:014101.