Biophysical Journal Volume 73 November 1997 2337–2346

Predicting the Structure of Apolipoprotein A-I in Reconstituted High-Density Lipoprotein Disks

James C. Phillips,* Willy Wriggers,* Zhigang Li,* Ana Jonas,# and Klaus Schulten*

*Theoretical Biophysics, Beckman Institute and Department of Physics, and #Department of Biochemistry, College of Medicine at Urbana-Champaign, University of Illinois, Urbana, Illinois 61801 USA

ABSTRACT In reconstituted high-density lipoproteins, apolipoprotein A-I and phosphatidylcholines combine to form disks in which the amphipathic α-helices of apolipoprotein A-1 bind to the edge of a lipid bilayer core, shielding the hydrophic lipid tails from the aqueous environment. We have employed experimental data, sequence analysis, and molecular modeling to construct an atomic model of such a reconstituted high-density lipoprotein disk consisting of two apolipoprotein A-I proteins and 160 palmitoyloleoylphosphatidylcholine lipids. The initial globular domain (1–47) of apolipoprotein A-I was excluded from the model, which was hydrated with an 8-Å shell of water molecules. Molecular dynamics and simulated annealing were used to test the stability of the model. Both head-to-tail and head-to-head forms of a reconstituted high-density lipoprotein were simulated. In our simulations the protein contained and adhered to the lipid bilayer while providing good coverage of the lipid tails.

INTRODUCTION

Lipoproteins are soluble complexes of proteins and lipids that transport lipids in the blood of vertebrates. Several classes of lipoproteins, characterized by their density, size, and protein composition, shuttle triglycerides and cholesterol from the intestine or liver to various tissues for storage or utilization, or transport cholesterol from the periphery to the liver for excretion or recycling. The latter function is performed by the high-density lipoproteins (HDLs). These lipoproteins are the smallest (~10 nm) and most dense class, composed of a 28-kDa protein called apolipoprotein A-I (apoA-I), several minor proteins, phospholipids, cholesterol, and cholesterol esters.

Apolipoprotein A-I defines the size and shape of HDL, solubilizes its lipid components, removes cholesterol from peripheral cells, activates the reaction of lecithin cholesterol acyltransferase (LCAT) that converts cholesterol to cholesterol esters in circulation, and delivers the cholesterol esters to the liver or steroidogenic tissues via cell surface receptors (Pownall and Gotto, 1992). Because of the importance of its functions, there is great interest in the structure of apoA-I, particularly in its lipid-bound state. So far, the structure of apoA-I in native lipoproteins has not been amenable to study, as it exists in very heterogeneous particles and, possibly, in diverse conformations. Reconstituted HDL (rHDL), of defined compositions and sizes, has provided the best opportunities to learn about the structure-function relationships of apoA-I (Jonas, 1986). Apolipoprotein A-I in rHDL complexes with phosphatidylcholines forms discoidal structures with a phospholipid bilayer core surrounded by a protein shell made up largely of amphipathic α-helices. The diameters of the disks vary, depending on the number of apoA-I molecules per particle (two, three, or four) and on the conformation of individual apoA-I molecules that may contain from six to eight amphipathic helices (Wald et al., 1990b; Jonas et al., 1989). These rHDLs are analogous in structure to nascent, native HDLs and exhibit all of the known physiological functions of HDLs. Thus they are excellent models for structural studies of apoA-I in the lipid-bound state.

A consensus model has emerged for the organization of apoA-I in discoidal rHDL particles from a variety of experimental approaches. Analysis of the gene structure of apoA-I has shown that it consists of four exons, and that exons 3 and 4 encode the mature protein. Starting at exon 4 (residue 44), the apoA-I sequence is characterized by homologous repeats of 11 or 22 amino acids, punctuated by Pro residues (Boguski et al., 1986). Secondary structure predictions, based on the primary sequence of apoA-I, suggest the presence of seven to nine amphipathic helices that coincide with the 22 amino acid repeats (Segrest et al., 1992). The dimensions of the rHDL disks and the 1.5-nm-wide protein shell (Wlodawer et al., 1979) suggest that the helices (~70% α-helix content from circular dichroism measurements) are arranged as tightly packed antiparallel segments running parallel to the lipid acyl chains of the bilayer phospholipids (Jonas, 1986; Wald et al., 1990b). The almost parallel orientation of the α-helices to the lipid acyl chains is supported by polarized attenuated total reflection–Fourier transform infrared spectroscopy measurements (Wald et al., 1990a). Finally, the pattern of monoclonal antibody binding to specific apoA-I epitopes is also consistent with a linear, antiparallel helix organization of apoA-I in rHDL particles (Marcel et al., 1991).

Because of the large sizes and inherently heterogeneous and dynamic structures of HDLs and rHDLs, attempts to crystallize lipid-bound apoA-I in forms suitable for x-ray

Received for publication 20 May 1997 and in final form 13 August 1997.

Address reprint requests to Dr. Klaus Schulten, Department of Physics, Beckman Institute 3147, University of Illinois, 405 N. Mathews Ave., Urbana, IL 61801. Tel.: 217-244-1604; Fax: 217-244-6078; E-mail: kschulte@ks.uiuc.edu.

0006-3495/97/11/2337/10 $2.00

structure determination have not been successful. Crystallization attempts on lipid-free apoA-I have been hampered by the ill-defined structure of the protein in solution and its tendency to self-associate at high concentrations (Davidson et al., 1996). The recent crystallization of a large fragment of apoA-I, lacking the N-terminal end (~20%) of the molecule, in a dimer form (Borhani et al., 1997) will provide much-needed information on the folding of apoA-I in solution. However, having a partial structure of apoA-I in solution will not obviate the need to elucidate its structure in a lipid-bound state, because it is well known that apoA-I free of lipid has a structure markedly different from those of the functional, lipid-bound forms of apoA-I (Davidson et al., 1996). In fact, the known X-ray structure of the apolipoprotein E N-terminal domain in water (Wilson et al., 1991) has not provided a clear understanding of the structure or function of this related apolipoprotein in its lipid-bound state.

Previous computer modeling by Brasseur and co-workers (Brasseur et al., 1990, 1992), based on the experimental data available on lipid-bound apoA-I, demonstrated that an antiparallel arrangement of amphipathic α-helices is energetically and sterically possible and suggested pairwise lateral electrostatic interactions between helices. The present work includes a similar analysis, but does not neglect the interhelical residues. In addition, the great increase in computer power since this earlier work was done has allowed our model to be tested in a molecular dynamics simulation of the full rHDL disk (protein, lipid, and water) in both head-to-head and head-to-tail configurations.

Although general protein structure prediction is not yet feasible, several properties of the rHDL system have motivated this attempt. Particularly encouraging is the form of the HDL particle in which apoA-I shields the lipid chains from the aqueous environment. Recent success in the prediction of a membrane protein (Hu et al., 1995) and other work (Engelman et al., 1986) suggests that predicting such structures is easier than predicting structures of globular proteins. In this work, we do not attempt to predict the structure of the N-terminal globular domain, but only of the lipid-bound region. Experimental evidence that apoA-I is primarily α-helical, combined with patterns of sequence repeats and proline punctuation, provides additional clues to secondary structure. Finally, the amphipathic nature of the predicted helices constrains the main elements of tertiary structure. This body of supporting evidence makes it likely that the predicted structure presented herein accurately represents the main features of apoA-I as found in HDL.

In the next section we describe the experimental constraints that influenced our model and the structure prediction and refinement methods employed. We then narrate the steps in the prediction process and present the final structures. Finally, we discuss the structure-function relationships implied by our model and the possibilities for future study.

METHODS

In this section we outline first the known experimental properties of apoA-I that allowed us to predict the structure of a dimer in complex with palmitoyloleoylphosphatidylcholine (POPC) lipids. Subsequently, we describe the methodology of the structure prediction and the molecular dynamics protocols used to refine the complex structure in solution. Note that, in the context of this paper, a "dimer" of apoA-I refers to two molecules of apoA-I that are bound to phospholipid in the discoidal rHDL complex; specific protein-protein interactions may or may not be present.

Physical properties of the apolipoprotein A-I–lipid system

The experimental parameters that have formed the basis for the present modeling of the apoA-I structure in rHDL complexes with POPC are listed in Table 1. Table 1 also gives the methods used in the determination of the various parameters and the appropriate literature references.

The rHDL particles have been prepared by the Na cholate reconstitution method (Matz and Jonas, 1982; Jonas, 1986), using purified human plasma apoA-I and commercially available (Sigma Chemical Co., St. Louis, Mo) palmitoyloleoylphosphatidylcholine (POPC). The resulting mixture of particles has been fractionated by gel filtration to give uniform and highly reproducible particles. Most of the experiments have been performed at 20–25°C. At this temperature, POPC exists in the liquid crystalline state within the particles, as shown by fluorescence polarization measurements.

TABLE 1 Experimentally determined parameters for apoA-I rHDL complexes with POPC

Structural parameter	Value	Experiment	References
ApoA-I molecules/particle	2	Chemical cross-linking and SDS-PAGE	Jonas et al., 1990; Davidson et al., 1994; McGuire et al., 1996
POPC molecules/particle	160 ± 8	Phosphate/protein analysis	Jonas et al., 1990; Davidson et al., 1994; McGuire et al., 1996
Particle diameter	9.8 ± 0.2 nm	Nondenaturing gradient gel electrophoresis	Jonas et al., 1990; Davidson et al., 1994; McGuire et al., 1996
	10.5 ± 1.9 nm	Electron microscopy	Nichols et al., 1984
Particle width	4.7 ± 0.6 nm	Electron microscopy	Nichols et al., 1984
α-helix content	74 ± 3%	CD spectroscopy	Jonas et al., 1990; McGuire et al., 1996

Except for the electron microscopy experiments (n = 1), all of the experiments have been reproduced on at least five independent preparations of rHDL. The mean values ± SD are given in the table.

The fluorescence polarization studies with lipophilic fluorescent probes have revealed that the POPC acyl chains are fluid, but somewhat restricted in motion relative to POPC in liposomes (Jonas et al., 1989).

Structure prediction

Exploiting the major constraint that the disklike bilayer of POPC lipids is surrounded by a dimer of mostly helical apoA-I, we used secondary structure prediction methods to identify the sequences corresponding to the "transmembrane" helices. This approach is analogous to the modeling techniques used in the structure prediction of integral membrane proteins (Johnson et al., 1994). In one aspect this problem differs from the prediction of integral membrane proteins: membrane-bound proteins typically involve apolar residues on the outside surface of the protein that are exposed to the hydrophobic lipid bilayer. In this case, one expects the protein to comprise amphiphilic helices that have their polar side exposed to the aqueous solution on the outside, and their nonpolar side exposed to the interior lipid phase.

To identify the helices, we employed the Holley/Karplus and Garnier/Osguthorpe/Robson (GOR) methods as implemented in Quanta 4.0 (MSI, 1994). The Holley/Karplus prediction (Holley and Karplus, 1989) identifies three types of secondary structure: helix, sheet, and coil. A neural network trained on a set of 48 representative protein structures assigns the secondary structure of a residue based on evidence of statistical correlation with secondary structure as far as eight places away in the sequence. GOR prediction (Garnier et al., 1978) identifies four secondary structure types: helix, extended chain, reverse turn, and coil. A residue is assigned the highest scoring conformational state based on a statistical analysis of 26 representative protein structures. Similar to the Holley/Karplus method, the GOR method computes the probability of each of the four conformational states from the information carried by amino acids within a 17-residue window centered at the prediction point.

After prediction of the eight transmembrane helices, we identified the orientation of each amphiphilic helix relative to the hydrophobic lipid phase by helical wheel projection (cf. Results). The helices were built with the program Quanta, version 4.0 (MSI, 1994), and positioned at a radius of 43 Å, with the hydrophobic side facing the inside. Thereby an experimentally determined outside diameter of 96 Å (Jonas, 1992) for the system was realized. The connecting loop regions and an external helix (7b) were modeled with Quanta.

The appropriate lipid for the construction of the membrane bilayer associated with apoA-I is POPC (Jonas, 1992). A POPC bilayer was created following the procedure described by Heller et al. (1993): a building block of four POPC lipids, created by Heller et al., was found suitable for duplication on a 5×5 grid to form one layer of the membrane. The resulting monolayer was duplicated, rotated, and shifted, resulting in a membrane patch of 200 lipids with a lateral area of 68 Å $\times$ 80 Å and a thickness (average phosphorus separation of the monolayers) of 44 Å (Heller et al., 1993). The apoA-I dimer encompasses 160 POPC lipids (Jonas et al., 1990; Davidson et al., 1994; McGuire et al., 1996). Hence we extracted an approximately disk-shaped patch of 160 lipids from the POPC bilayer of Heller et al. We created two apolipoprotein A-I dimers, one head-to-head and one head-to-tail, to investigate the dependence of the dimer stability on the unknown orientation of the proteins relative to each other. The 160-lipid bilayer disk was shrunk by 85%, compared to its geometry described above (Heller et al., 1993), to avoid steric clashes between the proteins and the lipids. Subsequently, the disk was fitted into the ringlike head-to-head and head-to-tail protein dimers. Both head-to-head and head-to-tail systems were energy-minimized to alleviate the structural effects of the manipulations necessary in the construction. The system size of the protein-lipid system was 27,850 atoms (two protein molecules of 3205 atoms each and 160 lipids of 134 atoms each).

Refinement

For refinement in aqueous solution, both protein-lipid systems were immersed in a shell of explicit water molecules of 8-Å thickness, which corresponds to approximately three layers of water molecules, following an X-PLOR procedure described by Brünger (1992). The water shell was constructed by covering the system with a three-dimensional lattice of water cubes that were provided with X-PLOR (Brünger, 1992). The cubes, of $(15.55\ \text{Å})^3$ size each, contained 125 water molecules, resulting in uniformly distributed water of density 1.0 g H_2O/cm^3. Water molecules inside a cylinder of radius 40 Å and thickness 32 Å, aligned to the orientation of the disklike lipid bilayer, were deleted. Thereby, a purely hydrophobic environment was realized in the hydrocarbon phase of the lipid bilayer, and a dry lipid-protein interface was provided for the hydrophobic protein-lipid interactions. After solvation, the total size of the head-head system was 46,222 atoms (6124 water molecules), and the size of the head-tail system was 46,522 atoms (6224 water molecules).

Subsequently, both the head-head and head-tail complexes were refined by using simulated annealing simulation protocols. Simulated annealing is a conformational search technique (van Laarhoven and Aarts, 1987; Brünger et al., 1990; Brünger, 1991) that allows biomolecules to pass barriers between local minima in the rugged potential energy landscape (Frauenfelder and Sligar, 1991) faster than would be realized by simulations at physiological temperature. The method thereby accelerates the relaxation of the system from any artifactual misfoldings created by the comparative modeling approach. In this work we used a maximum annealing temperature of 500 K; a higher temperature may adversely affect the stability of the protein structure (Xu et al., 1995). The molecular dynamics parameters of the CHARMM all-atom force field (Brooks et al., 1983), version 22, were used for the simulation of protein and lipids. The TIP3(P) water model (Jorgensen et al., 1983) was modified by omitting internal geometry constraints to provide water flexibility. This modification was motivated by MD studies that demonstrate an improvement of the physical properties of flexible water over the rigid TIP3 model (Teeter, 1991; Daggett and Levitt, 1993; Steinbach and Brooks, 1993). The simulations described in this paper were carried out with the program NAMD, version 2 (Nelson et al., 1996). A dielectric constant $\epsilon = 1$ and an integration time step of 1 fs were chosen. The r-RESPA integrator (Tuckerman et al., 1992) was employed so that nonbonded interactions could be calculated every other time step (Watanabe and Karplus, 1995). After an initial energy minimization, both systems were assigned initial velocities according to a Maxwell distribution and heated up to 500 K in steps of 25 K by velocity reassignments over a 20-ps time period. A cut-off distance of 10 Å for nonbonded interactions was chosen. The systems were then cooled down to 200 K in "fast" (35 ps) annealing protocols by using a 10-Å nonbonded cut-off distance. To investigate the convergence and stability of the refined structures, we carried out additional "slow" (70 ps) annealing protocols, involving a 12-Å nonbonded cut-off distance. The final structures were energy-minimized.

RESULTS

We now describe the intermediate conclusions in the secondary, tertiary, and quaternary structure prediction process. We then present and describe the properties of the predicted structures.

Structure prediction process

ApoA-I is believed to consist of an initial globular domain, followed by a region of mostly helical content. The first 43 residues correspond to exon 3 of the apoA-I gene (Boguski et al., 1986); exons are sometimes associated with distinct domains of a protein. Moreover, the fluorescence properties of the Trp residues of apoA-I, located in the N-terminal half of the molecule, indicate that they are protected from water whether apoA-I is free in solution or lipid-bound (Davidson et al., 1996). Furthermore, monoclonal antibody-binding

studies (Marcel et al., 1991) indicate that the N-terminal third of apoA-I has several overlapping discontinuous epitopes, suggesting the presence of tertiary structure. Finally, recent results on a deletion mutant of apoA-I that lacks the N-terminal 43 amino acids show that the mutant has the same lipid-binding properties as the intact protein and forms identical rHDL complexes with POPC (Rogers et al., 1997). Because the structure prediction of globular proteins is not yet reliable, we have neglected residues 1–47, attempting to predict the structure only of the helical region proposed for residues 48–243. This region is believed to form eight "transmembrane" helices in rHDL disks containing two apoA-I and a maximum number of lipids (Jonas et al., 1989).

The first step in structure prediction was to determine the locations of these helices. Fortunately, regularities in the sequence of apoA-I (Baker et al., 1975; Boguski et al., 1986), including the regular placement of proline residues, present clear evidence for helix boundaries. When arranged in pairs of undecapeptides as illustrated in Fig. 1, helix boundaries are marked by prolines at positions 66, 99, 121, 143, 165, 209, and 220. Although residue 187 is not a proline, the regular pattern of the sequence makes it likely that this is also a boundary in the eight-helix conformation. The segment 210–219 is not long enough to form a "transmembrane" helix, and we surmise that it may form a helix parallel to the plane of the membrane, connecting to the final "transmembrane" helix.

To strengthen these hypotheses, we analyzed the sequence with the Holley/Karplus and GOR secondary structure prediction algorithms. The scores for helix propensity are shown in Fig. 2, compared with helical regions in the final predicted structure. Results from both methods agree roughly with the inferences made from the sequence above, but in detail are somewhat conflicting. This was not unexpected, considering the unique nature of the rHDL particle and membrane proteins in general. For this reason, we relied mostly on proline placement and repeated patterns in the apoA-I sequence in determining helix location.

48	D N W D S V T	S T F S K L R E Q L G	helix 1
66	P V T Q E F W D N L E	K E T E G L R Q E M S	helix 2
88		K D L E E V K A K V Q	
99	P Y L D D F Q K K W Q	E E M E L Y R Q K V E	helix 3
121	P L R A E L Q E G A R	Q K L H E L Q E K L S	helix 4
143	P L G E E M R D R A R	A H V D A L R T H L A	helix 5
165	P Y S D E L R Q R L A	A R L E A L K E N G G	helix 6
187	A R L A E Y H A K A T	E H L S T L S E K A K	helix 7
209	P A L E D L R Q G L L		helix 7b
220		P V L E S F K V S F L	helix 8
231	S A L E E Y T K K L N	T Q	

FIGURE 1 Repeated sequences in apoA-I, adapted from Boguski et al. (1986). The modeled residues (48–243) are arranged in pairs of undecapeptides that show intrasequence similarity. Proline residues act as helix breakers, providing obvious boundaries between helical segments. Residues are colored by properties: acids are red, bases are blue, polar residues are green, hydrophobic residues are black, and prolines are yellow. Approximate predicted helix locations are 1:50–64, 2:70–87, 3:102–119, 4:124–139, 5:147–162, 6:171–183, 7:189–208, 7b:212–217, and 8:222–242.

The Holley/Karplus prediction coincides well with the suggested locations of helices 3–6, whereas 2, 7, and 8 are present, but less obvious. There is also a strong peak at the proposed connecting helix 7b between transmembrane helices 7 and 8. The proposed location of helix 1 shows a low helix propensity, in conflict with the pattern suggested by the sequence repeats. The GOR prediction agrees well with helices 3, 4, and 6, and to a lesser extent with the remaining transmembrane helices. Helix 7b, although suggested by Holley/Karplus and the location of proline residues, is not suggested by this method.

Given this secondary structure prediction, the next step was to arrange the helices into a plausible tertiary structure. This is a far less sensitive step with apoA-I than with tightly packed globular or membrane proteins, because apoA-I has far fewer protein-protein contacts and no rigid tertiary structure. In addition, the exact configuration of the turns is not crucial to the overall structure of the model. Thus we were not so concerned with exactly modeling the interhelical turns and the exact placement of helices relative to the lipids as with demonstrating a stable tertiary structure.

Fig. 3 shows the eight proposed transmembrane helices represented as helical wheels and shaded according to the Kyte and Doolittle hydrophobicity scale (Kyte and Doolittle, 1982). In this representation the amphipathic character of the helices becomes evident, and the helices can be readily arranged in an alternating up-down manner along a semicircle. With its amphipathic helices so aligned, the protein can be expected to adhere to the hydrophobic edge of the lipid disk.

To model a complete rHDL disk consisting of two apoA-I molecules, two copies of the predicted apoA-I structure were placed around a circular patch of POPC membrane as described above. There are two relative arrangements of these molecules: head-to-tail and head-to-head. In the head-to-tail configuration, the N terminus of one protein is adjacent to the C terminus of the other. Helices 1 and 8 from separate proteins are adjacent but aligned antiparallel, and the rHDL particle has a "top" and a "bottom" side. In the alternative head-to-head configuration, the 1 helices are adjacent and aligned parallel, as are the 8 helices.

The existence of rHDL disks containing three apoA-I molecules (Jonas et al., 1989) demonstrates the existence of the head-to-tail configuration, because such a complex contains odd numbers of heads and tails. This does not, however, preclude the existence of the head-to-head configuration or even its predominance. To investigate this alternative structure as well, systems with both configurations were constructed. Both models were solvated, and simulated annealing was conducted as described above to refine and test the stability of the structures. After minimization, the final structures presented below were obtained.

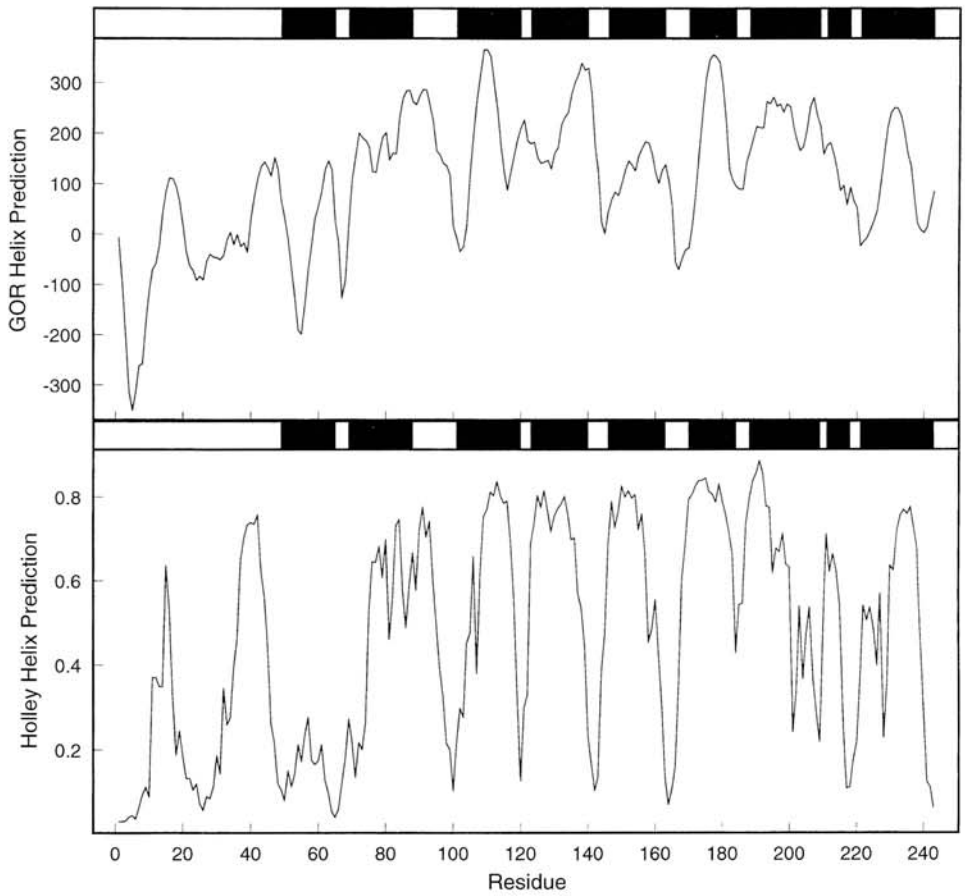

FIGURE 2 Helix propensity scores from GOR (*top*) and Holley/Karplus (*bottom*) secondary structure prediction methods, compared with predicted helical segments (*black bands*). Other propensity scores have been omitted for clarity. The region of our predicted helix 1 received equally low helix and sheet scores from the Holley/Karplus method.

Predicted structures

We present results from the two longer simulated annealing runs; results from the shorter runs were similar. The head-to-tail and head-to-head dimers combined yield a sample of four apoA-I molecules for analysis. Because of the fluidity of the POPC disk and the flexibility of the protein itself, the structures of the four molecules diverged during the simulation to illustrate a range of similar configurations rather than converging on a single ideal structure. Of course, these simulations are too short for any large rearrangement of the protein structure to occur and, hence, the final structures are strongly correlated with the original predicted structure.

Fig. 4 presents our final predicted structure for apoA-I as found in rHDL disks. All of the predicted helices are intact, although insufficient modeling of the interhelical regions has caused some disorder. The transmembrane helices are perpendicular to the bilayer, with the exception of helix 6, which has a slight tilt. Helix 6 is the shortest helix and is therefore distorted by tension from its connections to helices 5 and 7. In future modeling this may be eliminated by placing helices 5, 6, and 7 closer together initially. Other features of interest include the long, floppy region joining helices 2 and 3 and the short external helix 7b, which joins helices 7 and 8.

In the top image of Fig. 5 we see a side view of the rHDL model, which illustrates the manner and degree of coverage of the hydrophobic lipid tails. This is the region of least effective coverage, yet we see that water is excluded from ~80% of the tail area. Note that the stability of the rHDL disk does not depend on the ability of apoA-I to prevent the lipids from escaping, but rather on apoA-I stabilizing the interface between the lipid tails and the solvent. Hence this nontotal coverage is reasonable.

The bottom image of Fig. 5 details one of the two identical interfaces between apoA-I molecules in the head-to-tail configuration. Helices 1 and 8 are antiparallel and appear to be bound together. However, local interactions between the proteins are one-tenth the size of the affinity of apoA-I for the lipid bilayer, and hence, protein-lipid interaction is the primary factor in stabilizing the complex. One should bear in mind that the initial globular domain is missing from this model, and as such the detailed interactions between helices 1 and 8 cannot be characterized.

The missing globular domain also limits the inferences that can be made from the interface between helices 1 shown in the top image of Fig. 6. Neither a large interaction area or attraction are apparent. This is not the case, however, for the interface between helices 8 shown in the bottom image. This long helix presents a large surface area for interaction. In addition, the residues modeled at present as the external helix 7b could be arranged to create more complex tail-tail interactions, such as the exchange of helix 8. Hence our model supports the possible existence of the head-head configuration.

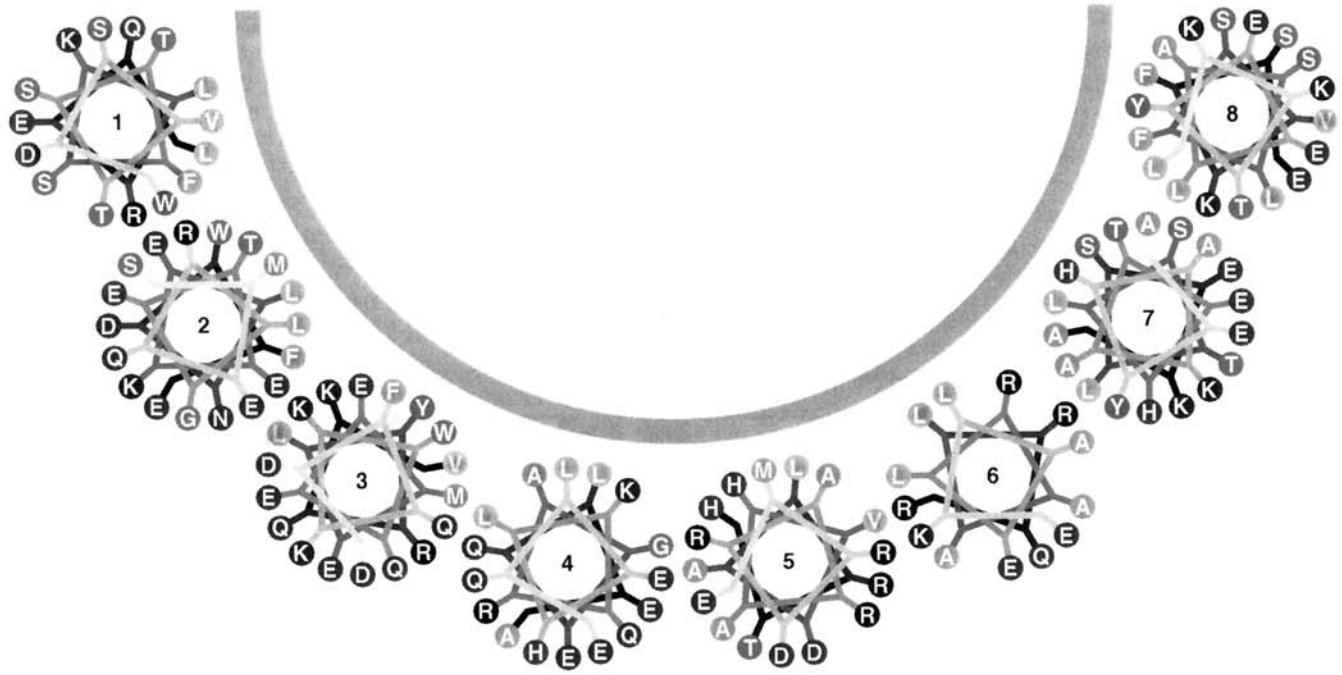

FIGURE 3 Schematic of helix amphipathic character, placement, and orientation in apoA-I. Residues are colored according to the hydrophobicity scale of Kyte and Doolittle (1982), with lighter shades more hydrophobic. Residues represented by wheels are 1:50–64, 2:70–87, 3:102–119, 4:124–139, 5:147–162, 6:171–183, 7:190–207, and 8:223–240. Odd-numbered helices are descending, even-numbered helices are ascending. This figure is not meant to represent helix spacing, as the helix representations have been enlarged for visibility.

The experimentally determined diameter and lipid content of the model are such that the protein cannot have close contacts between all pairs of adjacent helices. In additional energetic calculations, all pairs of adjacent helices had favorable net electrostatic interactions on the order of 25 to 100 kcal/mol, with the exception of 6/7 and 7/8, which were slightly unfavorable (~10 kcal/mol). This suggests greater uncertainty regarding the structure of helix 7, which is reasonable given the adjacent helix 7b. Regardless, this is a fraction of the binding energy between the helices and the lipid bilayer.

Figs. 7 and 8 present Ramachandran plots for our predicted structures. Eighty-six percent of residues are found in the most-favored regions of (ϕ, ψ) space. Some residues were found in the disallowed regions, particularly near (60°, −90°). All of these poorly configured residues are found in the loop regions between transmembrane helices, indicating the need for additional modeling in these areas. The exact configuration of the turns, however, is not crucial to the overall structure of the model, with eight helices binding perpendicular to the plane of the rHDL disk. No residues from the more carefully modeled external helix 7b region

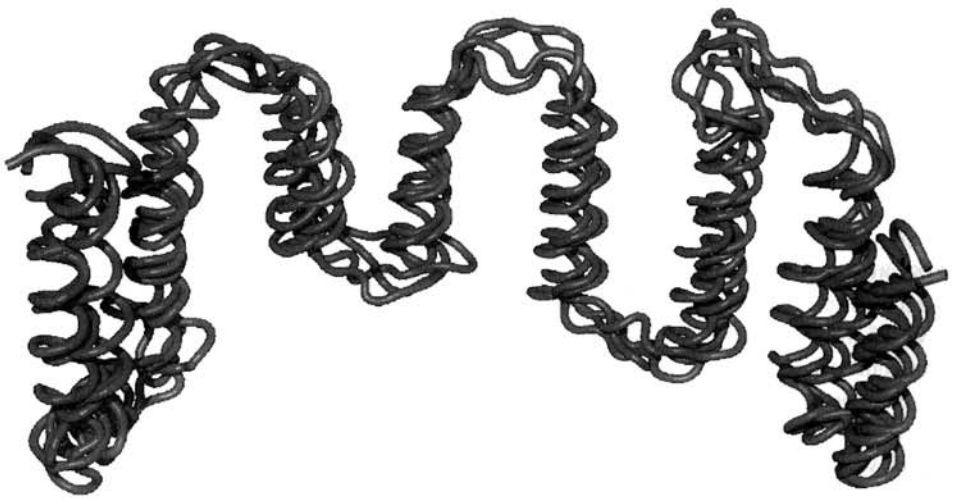

FIGURE 4 Final predicted structure of apoA-I. The molecule is oriented such that the top is the same as in Fig. 3 and helix 1 is at the far right. To emphasize the variability in the structure, the four apoA-I molecules from the two annealed dimers are aligned. There are no significant conformational differences in apoA-I between the head-to-tail and head-to-head models. Created with VMD (Humphrey et al., 1996) and Raster3D (Merritt and Murphy, 1994).

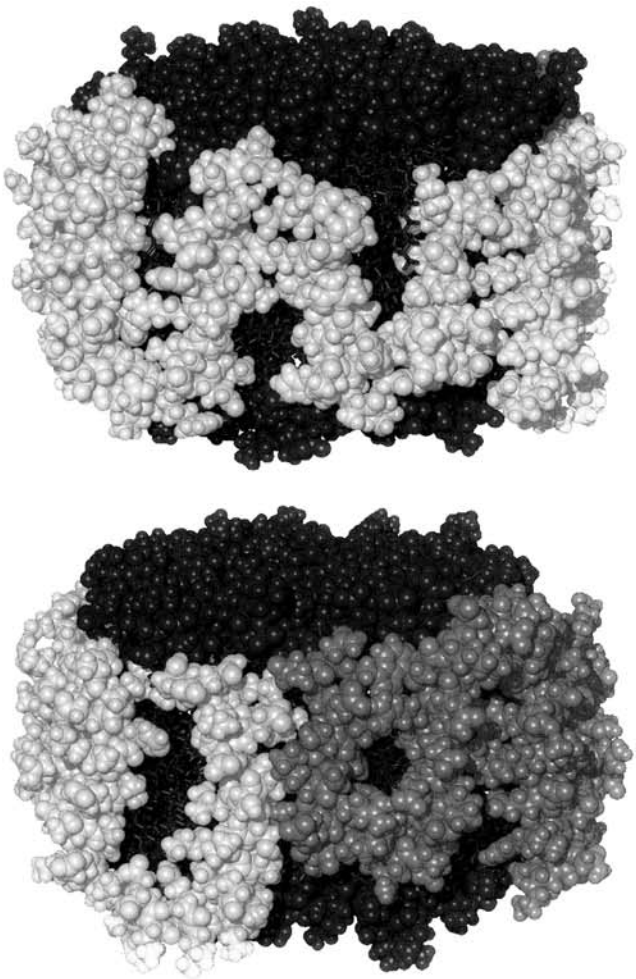

FIGURE 5 Final predicted structure of apoA-I in head-to-tail configuration. (*Top*) Side view of HDL particle, showing covering of lipid tails. The head region is at left; helices 4, 5, and 6 are centered; and the tail is at right. (*Bottom*) Interface between head (N-terminus) and tail (C-terminus) regions of the apoA-I proteins. The protein on the left exhibits its tail, and the protein on the right exhibits its head. Created with VMD (Humphrey et al., 1996) and Raster3D (Merritt and Murphy, 1994).

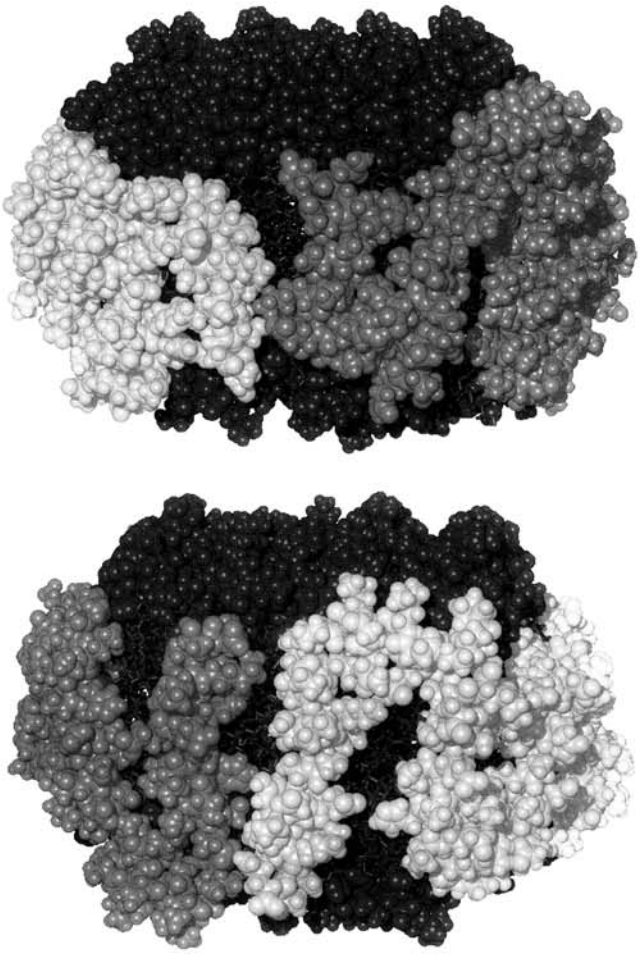

FIGURE 6 Final predicted structure of apoA-I in head-to-head configuration. (*Top*) Interface between heads (N-termini) of the apoA-I proteins. (*Bottom*) Interface between tails (C-termini) of the apoA-I proteins. Created with VMD (Humphrey et al., 1996) and Raster3D (Merritt and Murphy, 1994).

were in disallowed regions of the graph, suggesting a higher fidelity for the tail section of the model.

DISCUSSION

The most important feature of the present computer model of the apoA-I/POPC rHDL particle in water is the closely packed antiparallel array of amphipathic α-helices, providing good coverage of the phospholipid acyl chains. The proximity of adjacent helices suggests the possibility of further stabilizing interactions, in addition to the hydrophobic interactions between the residues on the nonpolar faces of the amphipathic helices and the phospholipid acyl chains. However, the size of the rHDL particle and the assumed motif of eight transmembrane helices preclude both strong interhelical interactions and contributions to lipid binding from the loop regions of the model structure.

Given the predominance of protein-lipid interactions over protein-protein interactions and the unknown contribution of the missing residues 1–47, the model cannot discriminate between the head-to-head and head-to-tail configurations. Two reports that address this question experimentally do not provide a definitive answer. A study of the competition of monoclonal antibodies binding to the rHDLs suggests that the two apoA-I molecules are organized in a head-to-tail configuration (Bergeron et al., 1995). However, the apoA-I–Milano dimer, which is constrained by a disulfide bridge at the mutated residues, Cys^{173}, into a head-to-head configuration, can readily form rHDL complexes with phospholipids (Calabresi et al., 1997). Therefore it is possible that both configurations exist.

On the basis of mutagenesis studies, as well as proteolytic digestion experiments, the C-terminal helices spanning residues 187–241 of apoA-I have been implicated in the initial binding of apoA-I to phospholipid bilayers and to HDL surfaces and in the interaction of apoA-I with cells (Davidson et al., 1996; Minnich et al., 1992; Luchoomun et al.,

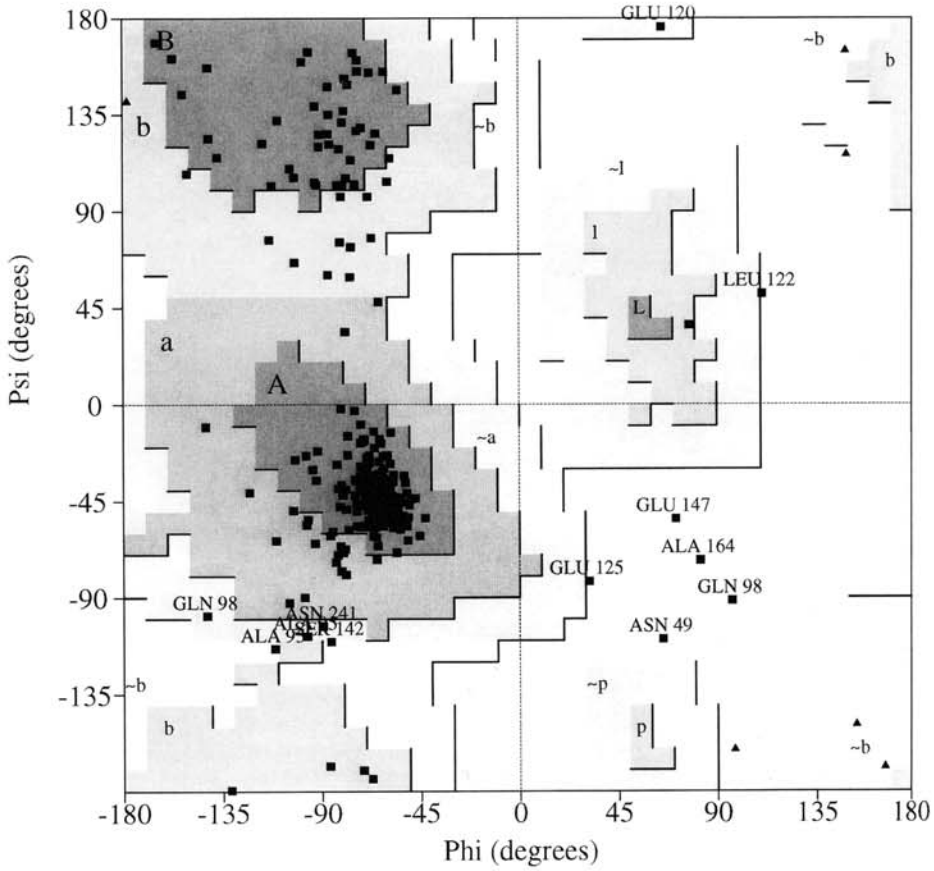

FIGURE 7 Ramachandran plot for final predicted structure of apoA-I in head-to-tail configuration. Created with PROCHECK (Laskowski et al., 1993).

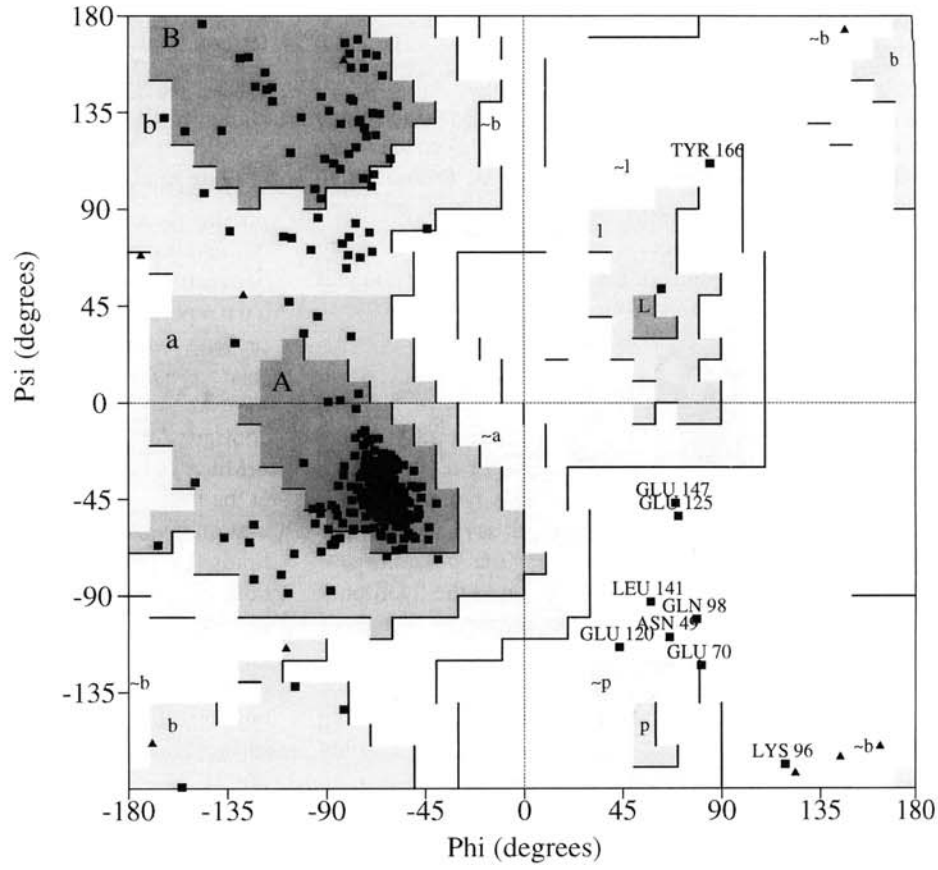

FIGURE 8 Ramachandran plot for final predicted structure of apoA-I in head-to-head configuration. Created with PROCHECK (Laskowski et al., 1993).

1994). This region of apoA-I is also accessible to proteases in rHDL particles, which are otherwise very resistant to proteolysis (Ji and Jonas, 1995). In fact, the short helix 7b that joins helices 7 and 8 in the model and appears to protrude from the edge of the particle could represent an accessible region of apoA-I.

Helices 3–4 and/or 5–6 have been implicated in the activation of the LCAT reaction (Sorci-Thomas et al., 1993; Holvoet et al., 1995). Furthermore, monoclonal antibody binding studies have suggested that helices 3–4 can swing out from the edge of the disk to permit the formation of smaller discoidal rHDLs when phospholipids are depleted (Marcel et al., 1991). Structural rearrangements involved in LCAT activation and particle transformations are clearly not precluded by this model, and could be tested in future modeling.

Future modeling may also address the transformation from a discoidal to a spherical rHDL particle, a process catalyzed by LCAT in vivo and in vitro (Jonas et al., 1990). Finally, the anticipated availability of a lipid-free crystal structure of apoA-I will provide secondary structure data that can be incorporated into an improved model of helix packing in rHDL.

This work was supported by the Roy J. Carver Charitable Trust, the National Institutes of Health (NIH PHS 5 P41 RR05969-04 and NIH HL 16059), and the National Science Foundation (NSF BIR 94-23827 EQ). JCP was supported by a Computational Science Graduate Fellowship from the United States Department of Energy.

REFERENCES

Baker, H. N., A. M. Gotto, Jr., and R. L. Jackson. 1975. The primary structure of human plasma high density apolipoprotein glutamine I (ApoA-I). *J. Biol. Chem.* 250:2725–2738.

Bergeron, J., P. G. Frank, D. Scales, Q.-H. Meng, G. Castro, and Y. L. Marcel. 1995. Apolipoprotein A-I conformation in reconstituted discoidal lipoproteins varying in phospholipid and cholesterol content. *J. Biol. Chem.* 270:27429–27438.

Boguski, M. S., M. Freeman, N. A. Elshourbagy, J. M. Taylor, and J. I. Gordon. 1986. On computer-assisted analysis of biological sequences: proline punctuation, consensus sequences, and apolipoprotein repeats. *J. Lipid. Res.* 27:1011–1034.

Borhani, D. W., D. P. Rogers, J. A. Engler, and C. G. Brouillette. 1997. Crystal structure of truncated human apolipoprotein A-I suggests a lipid-bound conformation. *Proc. Natl. Acad. Sci. USA.* In press.

Brasseur, R., L. Lins, B. Vanloo, J.-M. Ruysschaert, and M. Rosseneu. 1992. Molecular modeling of the amphipathic helices of the plasma apolipoproteins. *Proteins Struct. Funct. Genet.* 13:246–257.

Brasseur, R., J. D. Meutter, B. Vanloo, E. Goormaghtigh, J.-M. Ruysschaert, and M. Rosseneu. 1990. Mode of assembly of amphipathic helical segments in model high-density lipoproteins. *Biochim. Biophys. Acta.* 1043:245–252.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMm: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Brünger, A. 1991. Simulated annealing in crystallography. *Annu. Rev. Phys. Chem.* 42:197–223.

Brünger, A. T. 1992. X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR. The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT.

Brünger, A. T., A. Krukowski, and J. Erickson. 1990. Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr.* A46:585–593.

Calabresi, L., G. Franceschini, A. Burkybile, and A. Jonas. 1997. Activation of lecithin cholesterol acyltransferase by a disulfide-linked apolipoprotein A-I dimer. *Biochem. Biophys. Res. Commun.* 232:345–349.

Daggett, V., and M. Levitt. 1993. Realistic simulations of native-protein dynamics in solution and beyond. *Annu. Rev. Biophys. Biomol. Struct.* 22:353–380.

Davidson, W. S., T. Hazlett, W. W. Mantulin, and A. Jonas. 1996. The role of apolipoprotein A-I domains in lipid binding. *Proc. Natl. Acad. Sci. USA.* 93:13605–13610.

Davidson, W. S., D. L. Sparks, S. Lund-Katz, and M. C. Phillips. 1994. The molecular basis for the difference in charge between pre-β- and α-migrating high density lipoproteins. *J. Biol. Chem.* 269:8959–8965.

Engelman, D. M., T. A. Steitz, and A. Goldman. 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* 15:321–353.

Frauenfelder, H., and S. G. Sligar. 1991. The energy landscapes and motions of proteins. *Science.* 254:1598–1603.

Garnier, J., D. Osguthorpe, and B. Robson. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120:97–120.

Heller, H., M. Schaefer, and K. Schulten. 1993. Molecular dynamics simulation of a bilayer of 200 lipids in the gel and in the liquid crystal phases. *J. Phys. Chem.* 97:8343–8360.

Holley, L. H., and M. Karplus. 1989. Protein secondary structure prediction with a neural network. *Proc. Natl. Acad. Sci. USA.* 86:152–156.

Holvoet, P., Z. Zhao, B. Vanloo, R. Vos, E. Deridder, A. Dhoest, J. Taveirne, E. Brouwers, E. Demarsin, Y. Engelborghs, M. Rosseneu, D. Collen, and R. Brasseur. 1995. Phospholipid binding and lecithin-cholesterol acyltransferase activation properties of apolipoprotein A-I mutants. *Biochemistry.* 34:13334–13342.

Hu, X., D. Xu, K. Hamer, K. Schulten, J. Köpke, and H. Michel. 1995. Predicting the structure of the light-harvesting complex II of *Rhodospirillum molischianum*. *Protein Sci.* 4:1670–1682.

Humphrey, W. F., A. Dalke, and K. Schulten. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* 14:33–38.

Ji, Y., and A. Jonas. 1995. Properties of an N-terminal proteolytic fragment of apolipoprotein A-I in solution and in reconstituted high density lipoproteins. *J. Biol. Chem.* 270:11290–11297.

Johnson, M. S., N. Srinivasan, R. Sowdhamini, and T. L. Blundell. 1994. Knowledge-based protein modeling. *Crit. Rev. Biochem. Mol. Biol.* 29:1–68.

Jonas, A. 1986. Reconstitution of high-density lipoproteins. *Methods Enzymol.* 128:553–582.

Jonas, A. 1992. Lipid-binding properties of apolipoproteins. *In* Structure and Function of Apolipoproteins. M. Rosseneu, editor. CRC Press, Boca Raton, FL. 217–250.

Jonas, A., K. E. Kezdy, and J. H. Wald. 1989. Defined apolipoprotein A-I conformations in reconstituted high density lipoprotein discs. *J. Biol. Chem.* 264:4818–4824.

Jonas, A., J. H. Wald, K. L. H. Toohill, E. S. Krul, and K. E. Kezdy. 1990. Apolipoprotein A-I structure and lipid properties in homogeneous reconstituted spherical and discoidal high density lipoproteins. *J. Biol. Chem.* 265:22123–22129.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105–132.

Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.

Luchoomun, J., N. Theret, V. Clavey, P. Duchateau, M. Rosseneu, R. Brasseur, P. Denefle, J.-C. Fruchart, and G. R. Castro. 1994. Structural domain of apolipoprotein A-I involved in its interaction with cells. *Biochim. Biophys. Acta.* 1212:319–326.

Marcel, Y. L., P. R. Prevost, H. Kao, E. Raffai, N. VuDac, J.-C. Fruchart, and E. Rassart. 1991. The epitopes of apolipoprotein A-I define distinct

structural domains including a mobile middle region. *J. Biol. Chem.* 266:3644–3653.

Matz, C. E., and A. Jonas. 1982. Micellar complexes of human apolipoprotein A-I with phosphatidylcholines and cholesterol prepared from cholate-lipid dispersions. *J. Biol. Chem.* 257:4535–4540.

McGuire, K. A., W. S. Davidson, and A. Jonas. 1996. High yield overexpression and characterization of human recombinant proapolipoprotein A-I. *J. Lipid. Res.* 37:1519–1528.

Merritt, E. A., and M. E. P. Murphy. 1994. Raster3D version 2.0—a program for photorealistic molecular graphics. *Acta Crystallogr. D.* 50: 869–873.

Minnich, A., X. Collet, A. Roghani, C. Cladaras, R. L. Hamilton, C. J. Fielding, and V. I. Zannis. 1992. Site-directed mutagenesis and structure-function analysis of the human apolipoprotein A-I. *J. Biol. Chem.* 267:16553–16560.

MSI. 1994. QUANTA 4.0. Molecular Simulations, Burlington, MA.

Nelson, M., W. Humphrey, A. Gursoy, A. Dalke, L. Kalé, R. D. Skeel, and K. Schulten. 1996. NAMD—a parallel, object-oriented molecular dynamics program. *J. Supercomput. Appl.* 10:251–268.

Nichols, A. V., E. L. Gong, P. J. Blanche, T. M. Forte, and V. G. Shore. 1984. Interaction of model discoidal complexes of phosphatidylcholine and apolipoprotein A-I plasma components. *Biochim. Biophys. Acta.* 793:325–337.

Pownall, H. J., and A. M. Gotto, Jr. 1992. Human plasma apolipoproteins in biology and medicine. *In* Structure and Function of Apolipoproteins. M. Rosseneu, editor. CRC Press, Boca Raton, FL. 1–32.

Rogers, D. P., C. G. Brouillette, J. A. Engler, S. W. Tendian, L. Roberts, V. K. Mishra, G. M. Anantharamaiah, S. Lund-Katz, M. C. Phillips, and M. J. Ray. 1997. Truncation of the amino terminus of human apolipoprotein A-I substantially alters only the lipid-free conformation. *Biochemistry.* 36:288–300.

Segrest, J. P., M. K. Jones, H. DeLoof, C. G. Brouillette, Y. V. Venkatachalapathi, and G. M. Anantharamaiah. 1992. The amphipathic helix in the exchangeable apolipoproteins: a review of secondary structure and function. *J. Lipid. Res.* 33:141–166.

Sorci-Thomas, M., M. W. Kearns, and J. P. Lee. 1993. Apolipoprotein A-I domains involved in lecithin-cholesterol acyltransferase activation. *J. Biol. Chem.* 268:21403–21409.

Steinbach, P. J., and B. R. Brooks. 1993. Protein hydration elucidated by molecular dynamics simulation. *Proc. Natl. Acad. Sci. USA.* 90: 9135–9139.

Teeter, M. M. 1991. Water-protein interactions: theory and experiment. *Annu. Rev. Biophys. Biophys. Chem.* 20:577–600.

Tuckerman, M., B. J. Berne, and G. J. Martyna. 1992. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* 97:1990–2001.

van Laarhoven, P. J. M., and E. H. L. Aarts. 1987. Simulated Annealing: Theory and Application. D. Reidel, Dordrecht, the Netherlands.

Wald, J. H., E. Goormaghtigh, J. D. Meutter, J.-M. Ruysschaert, and A. Jonas. 1990a. Investigation of the lipid domains and apolipoprotein orientation in reconstituted high density lipoproteins by fluorescence and IR methods. *J. Biol. Chem.* 265:20044–20050.

Wald, J. H., E. S. Krul, and A. Jonas. 1990b. Structure of apolipoprotein A-I in three homogeneous reconstituted high density lipoprotein particles. *J. Biol. Chem.* 265:20037–20043.

Watanabe, M., and M. Karplus. 1995. Simulation of macromolecules by multiple-time-step methods. *J. Phys. Chem.* 99:5680–5697.

Wilson, C., M. R. Wardell, K. H. Weisgraber, R. W. Mahley, and D. A. Agard. 1991. Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E. *Science.* 252:1817–1822.

Wlodawer, A., J. P. Segrest, B. H. Chung, R. Chiovetti, Jr., and J. N. Weinstein. 1979. High-density lipoprotein recombinants: evidence for a bicycle tire micelle structure obtained by neutron scattering and electron microscopy. *FEBS Lett.* 104:231–235.

Xu, D., M. Sheves, and K. Schulten. 1995. Molecular dynamics study of the M412 intermediate of bacteriorhodopsin. *Biophys. J.* 69:2745–2760.