A Three-Dimensional Molecular Model of Lipid-Free Apolipoprotein A-I Determined by Cross-Linking/Mass Spectrometry and Sequence Threading†

R. A. Gangani D. Silva,‡ George M. Hilliard,§ Jianwen Fang,¶ Stephen Macha,# and W. Sean Davidson*,‡

Department of Pathology and Laboratory Medicine, University of Cincinnati, Cincinnati, Ohio 45267, Department of Molecular Sciences and Center of Excellence in Genomics and Bioinformatics, University of Tennessee Health Science Center, Memphis, Tennessee 38163, Bioinformatics Core Facility, University of Kansas, Lawrence, Kansas 66045, and Department of Chemistry, Mass Spectrometry Services, University of Cincinnati, Cincinnati, Ohio 45221

Received October 26, 2004; Revised Manuscript Received December 10, 2004

ABSTRACT: Apolipoprotein (apo) A-I, a 243-residue, 28.1-kDa protein is a major mediator of the reverse cholesterol transport (RCT) pathway, a process that may reduce the risk of cardiovascular disease in humans. In plasma, a small fraction of lipid-free or lipid-poor apoA-I is likely a key player in the first step of RCT. Therefore, a basic understanding of the structural details of lipid-free apoA-I will be useful for elucidating the molecular details of the pathway. To address this issue, we applied the combined approach of cross-linking chemistry and high-resolution mass spectrometry (MS) to obtain distance constraints within the protein structure. The 21 lysine residues within apoA-I were treated with homo bifunctional chemical cross-linkers capable of covalently bridging two lysine residues residing within a defined spacer arm length. After trypsin digestion of the sample, individual peptide masses were identified by MS just after liquid chromatographic separation. With respect to the linear amino acid sequence, we identified 5 short-range and 12 long-range cross-links within the monomeric form of lipid-free apoA-I. Using the cross-linker spacer arm length as a constraint for identified Lys pairs, a molecular model was built for the lipid-free apoA-I monomer based on homology with proteins of similar sequence and known three-dimensional structures. The result is the first detailed model of lipid-free apoA-I. It depicts a helical bundle structure in which the N- and C-termini are in close proximity. Furthermore, our data suggest that the self-association of lipid-free apoA-I occurs via C- and N-termini of the protein based on the locations of six cross-links that are unique to the cross-linked dimeric form of apoA-I.

High levels of high-density lipoprotein (HDL)[1] in plasma are associated with a decreased risk of cardiovascular disease (CVD) in humans (*1*). Apolipoprotein (apo) A-I, the major protein constituent of HDL, has been implicated as a critical mediator of many known functions of HDL. For example, apoA-I interacts with lecithin:cholesterol acyl transferase (LCAT) for HDL maturation (for a review, see ref *2*) and may act as a ligand for scavenger-receptor B1 (SR-B1) for delivery of cholesteryl esters from HDL to the liver for excretion and recycling (*3*). In addition, several studies (reviewed in ref *4*) have indicated that lipid-free/lipid-poor apoA-I plays a prominent role in the stimulation of cholesterol efflux from peripheral cells via the ATP-binding cassette transporter (ABCA1).

The molecular details of these processes are not yet fully understood, arguably due to a lack of understanding of the structural adaptations of lipid-free and lipid-bound apoA-I. Many of the structural studies on this single polypeptide, 243-residue, 28.1-kDa protein have been carried out in reconstituted discoidal HDL complexes containing two molecules of apoA-I per disc (reviewed in ref *5*). These homogeneous in vitro preparations have been used with great success to gain insights into the structure of lipid-associated apoA-I. Indeed, a molecularly detailed computer model has been proposed (*6*−*8*) that appears to be well supported by recent experimental data (*9*−*13*).

Over the past two decades, a significant amount of work has also been directed at understanding the three-dimensional (3-D) structure of lipid-free apoA-I. This form presents a formidable challenge to traditional high-resolution structural

† This work was supported by Grants HL67093 and HL62542 (W.S.D.) from the NIH, a postdoctoral fellowship from the Ohio Valley Affiliate of the American Heart Association (R.A.G.D.), and NIH P20RR16475 (J.F.).

* To whom correspondence should be addressed. Department of Pathology and Laboratory Medicine, University of Cincinnati, 2120 Galbraith Rd., Cincinnati, OH 45237-0507 USA. Telephone: (513) 558-3707. Fax: (513) 558-1312. E-mail: Sean.Davidson@UC.edu.

‡ Department of Pathology and Laboratory Medicine, University of Cincinnati.

§ University of Tennessee Health Science Center.

¶ University of Kansas.

Department of Chemistry, Mass Spectrometry Services, University of Cincinnati.

[1] Abbreviations: apoA-I, apolipoprotein A-I; BS^3, (bis)sulfosuccinimidyl suberate; DSP, dithiobis (succinimidylpropionate); DMSO, dimethyl sulfoxide; FPLC, fast performance liquid chromatography; EPR-electron paramagnetic resonance; HDL, high-density lipoprotein; HPLC, high performance liquid chromatography; λ-max, wavelength of maximum fluorescence; MOE, molecular operating environment; MS, mass spectrometry; MWCO, molecular weight cut off; NRMSD, normalized root-mean-square difference; TIC, total ion chromatogram; NMR, nuclear magnetic resonance; PBS, phosphate buffered saline; PDB, protein data bank; POPC, 1-palmitoyl 2-oleoyl phosphatidylcholine; RCT-reverse cholesterol transport; RP-HPLC, revered phase-HPLC.; SDS−PAGGE, sodium dodecyl sulfate−polyacrylamide gradient gel electrophoresis; UV−CD, ultraviolet−circular dichroism.

10.1021/bi047717+ CCC: $30.25
Published on Web 02/01/2005

techniques such as NMR or X-ray crystallography because apoA-I spontaneously self-associates into a heterogeneous array of oligomers in a concentration-dependent manner. Despite this, Borhani et al. successfully crystallized a fragment of apoA-I that lacked the N-terminal 43 amino acids (*14*). The structure at a resolution of 4 Å showed a tetrameric, horse-shoe-shaped assembly composed of apoA-I molecules arranged in extended, kinked α-helices. Because of the high-salt conditions used for crystallization, the missing N-terminus and the oligomerization, the applicability of this structure to native monomeric apoA-I has not been clear (*15*, *16*). On the other hand, high-resolution NMR spectroscopic data to date on apoA-I is limited to the fragment 142−187 which spans helices 6 and 7 (*17*).

Most of the information obtained on full-length lipid-free apoA-I under native conditions has come from comparatively lower resolution approaches such as fluorescence spectroscopy (*18*, *19*), analytical ultracentrifugation (*20*), limited proteolysis (*21*, *22*), monoclonal antibody binding (*23*), and thermodynamics of unfolding on both wild-type and mutated forms of apoA-I (*15*, *24*−*26*). A general consensus from the literature is that lipid-free apoA-I likely exists as a quasistable association of amphipathic α-helices. This motif has strong support from successful X-ray crystal and NMR structures of other apolipoproteins composed of similar amphipathic α-helices such as apoE (*27*) and apolipophorin III (*28*). Beyond this general description of a "loose helical bundle" however, the specifics of which regions of apoA-I exist as helices and the spatial organization of how the helices interact is a subject of speculation. Two general models have been proposed. Roberts et al. (*21*) proposed a four helical bundle with α-helical segments placed largely by primary sequence analysis with unfolded regions corresponding to sites of proteolytic sensitivity. Agree et al. (*18*) proposed a more elongated model based on fluorescence energy transfer distance measurements between four introduced probes and all four naturally occurring tryptophan residues present in the N-terminal half of apoA-I. In addition, analytical ultracentrifugation studies have suggested the coexistence of both a compact helical bundle and an extended form that might be an intermediate that participates in lipid-binding (*20*). These schematic models have proven useful for driving further experimentation on apoA-I structure. However, they are limited to the level of secondary structure assignment and general molecular dimensions because they lack spatial information required for understanding the 3-D structure. Further understanding of how lipid-free apoA-I transitions to the lipid-bound state requires specific knowledge of which regions within apoA-I actually interact within the 3-D structure.

In the current study, we combined cross-linking chemistry and high-resolution mass spectroscopy to experimentally determine the points of contact in native lipid-free apoA-I. We previously used the same technique to study the spatial organization of apoA-I molecules on the edge of discoidal HDL particles (*12*). This approach, combined with homology modeling, has been used to correctly identify the 3-D fold of bovine basic fibroblast growth factor as compared to X-ray crystallographic and NMR structural data (*29*). In addition, the approach has been successfully used to build low-resolution 3-D structures for proteins of unknown structure (*30*, *31*). In this study, we identified 17 cross-links within

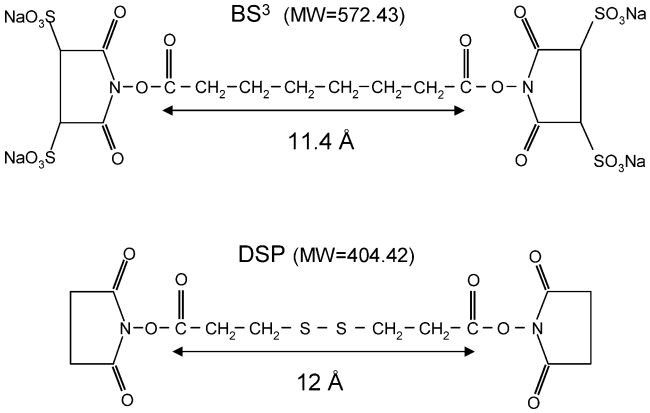

FIGURE 1: Chemical structures of bis(sulfosuccinimidyl) suberate (BS^3) and dithiobis (succinimidyl)propionate (DSP). The spacer arm length and the molecular weight of each cross-linker are given.

monomeric lipid-free apoA-I. This list of distance constraints was used to build a molecular model for apoA-I that fits much of the experimental data obtained to date on this form of apoA-I.

EXPERIMENTAL PROCEDURES

ApoA-I Cross-Linking and Trypsin Digestion. ApoA-I was isolated and purified from human plasma as described before (*12*). The protein was cross-linked at 1 mg/mL in phosphate buffered saline (PBS, pH 7.8) at 4 °C. The two cross-linkers used in these experiments, dithiobis (succinimidylpropionate) (DSP) and bis(sulfosuccinimidyl) suberate (BS^3) (Pierce, Rockford, IL) (Figure 1) were handled with minimum exposure to air to reduce hydrolysis. The cross-linker stock solutions were prepared to be 6.5 mg/mL (DSP in dimethyl sulfoxide [DMSO] and BS^3 in PBS) and added to the protein sample within 1 min of preparation. In DSP cross-linking experiments, the DMSO added was below 10% of the total volume of the protein solution. After cross-linker addition, the reaction solutions were incubated at 4 °C for 24 h, quenched with Tris buffer at a final concentration of 100 mM, dialyzed (3500 MWCO dialysis membrane) into 10 mM ammonium bicarbonate buffer (pH 8.1) to remove unbound cross-linker, and lyophilized. The cross-linked protein was solubilized in Tris buffer containing 3 M guanidine and the cross-linked protein components (internally cross-linked monomer, dimer, and higher-order oligomers) were separated in the same buffer on a gel filtration column set up (Superdex 200 and Superose 6 in tandem; Amersham Biosciences, USA) at 0.4 mL/min controlled by a Fast Performance Liquid Chromatography (FPLC) system (Amersham Biosciences, USA). Fractions corresponding to monomeric and dimeric cross-linked apoA-I were combined separately, dialyzed into ammonium bicarbonate buffer at 4 °C, concentrated by ultrafiltration (YM-10, MWCO 10,000, Millipore Corporation, Bradford, MA), and digested with sequencing grade trypsin (Promega, USA) at 5% w/w enzyme/protein at 37 °C for 2 h. Aliquots of the digested protein were lyophilized and stored at −20 °C until used for mass spectroscopic analysis.

Mass Spectrometry. Mass spectrometric measurements were performed on a Sciex QSTAR DE (equipped with an atmospheric electrospray ionizer and a quadrapole time-of-flight dual analyzer) connected to an on-line capillary HPLC (Agilent 1100). In brief, 150 pmol of the digested protein in

Table 1: Template Proteins Used for apoA-I Homology Modeling

protein	PDB ID	template residues[a]	represented apoA-I residues	sequence identity (%)
Ng Gtpase domain	1O87	37−87	15−57	31
apolipoprotein E4	1B68	54−162	58−166	24
apoA-I	1GW3	142−182	167−187	100
enoyl acyl carrier protein reductase	1ENP	137−178	200−241	33

[a] Note that templates of 1O87 and 1GW3 consist of more amino acids than the represented regions of apoA-I. The extra amino acids were used during the "merging step" of the intermediate models (see Experimental Procedures) and were removed subsequently.

0.1% trifluoroacetic acid (TFA) were injected into the HPLC and separated into component peptides by use of a C18 (Agilent ZORBAX SB 0.5 mm id × 15 cm) reversed phase column. The tryptic peptide elution was carried out by application of a complex acetonitrile gradient of 3−95% (3−20% by 10 min, 20−40% by 70 min, 40−50% by 80 min, and 50−95% by 90 min) at a flow rate of 7.5 (μL/min), which was optimized for the separation of tryptic peptides from human apoA-I. The eluting peaks were subjected to subsequent mass spectrometric detection in the range 100−2800 *m*/*z*. The instrument was calibrated internally and externally. The internal calibration was performed using a polyalanine calibration kit (200−2800 *m*/*z*, Sigma-Aldrich, USA) using manufacturer's instructions. The external calibration was performed using prominent peptide masses of apoA-I that cover the mass range of interest (780.4242 Da for residues 154−160, 1300.6412 Da for 161−171, and 1931.9265 Da for 62−77 (*12*)).

Data Analysis. The Analyst QS software (Applied Biosystems) was used to analyze the mass spectra. Individual mass spectra corresponding to each peak in the total ion chromatogram (TIC, the detector response proportional to the number of ions reaching the detector as a function of time) were generated and the detectable masses well above the noise level were recorded, indicating their charge states (ranged from +1 to +5). A monoisotopic mass list was generated by averaging all charge states for particular peptide masses. The software GPMAW (ChemSW, Inc. version 4.2) was used to identify the sequences of the unmodified and intrapeptide cross-links (see Results for our definitions of "intra- and interpeptide" cross-links). The program identifies the presence of hydrolyzed cross-links and other non-cross-link-related peptide modifications such as Met oxidation. Interpeptide cross-links were identified using a two-dimensional spread sheet that contains all theoretical masses of two tryptic peptides with an intervening cross-link (*12*).

Our criteria for identifying a cross-linked peptide mass was as follows: (1) The intensity of a given mass peak must be at least 5× background levels with at least two charge states present. (2) Putative identified peptides or cross-linked peptides must have a sufficient number of Lys residues available for modification that are not present at the extreme C-terminus (i.e., trypsin does not cleave after Lys residues modified by the cross-linker). (3) The peptides identified should not be partially cleaved by trypsin (i.e., the peptides should not contain skipped cleavable sites). (4) The experimental mass must match with the theoretical mass within ± 50 ppm. This was derived from repeated experiments measuring the maximal variation from peptides from untreated apoA-I. (5) A cross-linked peptide or peptides must be identified with both DSP and BS^3 cross-linkers. The presence of a specific cross-link was confirmed only if all of the above conditions were satisfied.

Homology Modeling. (a) General Strategy. The technique of homology modeling derives a hypothetical 3-D structure for a given protein sequence by statistically comparing it to a template sequence of known structure (by X-ray crystallography or NMR). Ideal templates for this analysis tend to share an evolutionary relationship, similar function, and greater than 30% sequence identity with the sequence being modeled. Although a limited number of exchangeable apolipoprotein structures have been solved, including regions of all three isoforms of human apolipoprotein E (*27*) and insect apolipophorins (*28*), homology modeling for apoA-I has not been possible because these proteins exhibit much less than 30% identity with apoA-I over its full length. To circumvent this problem, we employed a compartmentalized approach where sequences from a collection of template proteins with adequate homologies were used to build models for individual regions of apoA-I. The templates were chosen with regard to sequence homology, residue type, and amphipathic nature of the helices. An additional important criterion for template selection was the presence of apparent reverse turns in the protein backbone suggested by the experimentally derived cross-links. The individual models were then assembled in light of the cross-link distance constraints derived from this study and then subjected to an energy minimization algorithm to finish the model.

(b) Method Details. The Protein Data Bank (PDB) database was searched for protein sequences similar to human apoA-I using NCBI's online BLAST engine (http:// www.ncbi.nih.gov/BLAST/). The model building was performed using MOE (Molecular Operating Environment 2004.03, The Chemical Computing Group Inc., 2004) and AMBER 7. On the basis of the criteria mentioned above, residues 54−162 of apolipoprotein E4 (apoE4, PDB ID: 1B68) were used to model residues 58−166 of apoA-I (identities = 28/109 [24%], positives = 51/109 [47%]). Twenty-five intermediate models were created, each finely energy-minimized for steric interactions using AMBER-94 force field with solvation option activated. The best intermediate was then chosen for further consideration. Partial models were built using other the templates listed in Table 1 These were merged onto the existing intermediate and relative position of the helices were adjusted considering cross-link data and the amphipathic nature of the α-helices. Since we found no usable templates for the region 188−199 and there are disagreements in primary sequence structural prediction algorithms for the N-terminus 1−14, random structures were created for these segments. Two C-terminus residues 242−243 were also added as unordered conformations. The composite model was subjected to energy minimization in MOE with the cross-linking distance constraints. The resulting model was then minimized in explicit aqueous solution using the *Sander*

program of AMBER 7 package.[2] The protein was surrounded by a periodic truncated octahedron box of water molecules described by TIP3P potential (*32*) extended to a distance of 10 Å from any solute atom. The system was neutralized by sodium cations initially placed around the protein using a Coulombic potential on a grid. We performed the energy minimization in two stages. In the first stage, only water molecules were minimized while the protein was fixed. The second stage was to minimize the entire system as a whole. Each stage was conducted with 500 cycles of steepest descents followed by 4500 conjugate gradient minimization. Nonbonded cutoff was set to 12.0 Å. After two stages of minimization, the water molecules and sodium cations were stripped from the model using ptraj module of AMBER 7. Final minimization was done in MOE with the distance constraints applied and helix backbone atoms fixed.

Circular Dichroism (CD) Spectroscopy. The experimental apoA-I CD spectrum was acquired using a Jasco J-600 spectrometer with a 0.1-cm cuvette averaging four scans across a 180−260 nm range. The spectra were acquired with 30 μg/mL protein in 20 mM phosphate buffer at pH 7.8. The average secondary structural content of lipid-free apoA-I was determined by fitting the UV−CD spectrum of lipid-free apoA-I using different deconvolution algorithms provided by the Dicroweb service (http://www.cryst.bbk.ac.uk/cdweb/html/home.html). The best algorithm for fitting the experimental data was chosen by comparing a normalized root-mean-square difference (NRMSD), which reflects the average difference between the experimental data and a theoretical model generated by each algorithm based on a database of CD spectra from proteins of known structure. The CDSSTR (*33*, *34*) provided the best fit with an NRMSD of 0.02. In addition, far UV−CD spectra were also recorded (at room temperature) at higher concentrations (1 mg/mL), cross-linked at 4 °C for direct comparison with mass spectroscopic studies.

Fluorescence Spectroscopy. Fluorescence emission spectra of unmodified apoA-I and apoA-I cross-linked with BS^3 at a 1:15, protein: BS^3 molar ratio were recorded at both 4 °C and 25 °C. The cross-linking was carried out under conditions identical to those for the mass spectrometry samples described above. The fluorescence measurements were performed on a Photon Technology International Quantamaster spectrometer. The emission spectra of unmodified and cross-linked apoA-I were collected from 305 to 380 nm using Trp excitation wavelength of 295 nm chosen to minimize the Tyr fluorescence of apoA-I. The appropriate buffer controls were also included.

RESULTS

The Approach. This approach is based on the principle that cross-linking of apoA-I in solution with either BS^3 or DSP captures and locks the native conformation of the protein. The ϵ-amino groups of 21 lysine residues as well as the N-terminus serve as specific targets for the cross-linking agents (*12*, *29*, *35*). The formation of a cross-link is complete if both C═O functional groups of a cross-linker molecule (Figure 1) react with two free amino groups forming two covalent linkages. The cross-linked protein is then trypsinized into a series of peptides that may or may not be cross-linked to other peptides as dictated by the protein structure. The peptides are subjected to RP-HPLC separations followed by mass spectrometric detection with high mass measurement accuracy. In the current study, the use of two cross-linkers of similar spacer arm-lengths but considerably different masses provided a check on the identifications and virtually eliminated the possibility of a miss-assigned cross-link. For example, identical masses observed in both BS^3 and DSP cross-linking experiments corresponded to peptides unmodified by cross-linking (Figure 2a,b). On the other hand, a peptide that contained a single cross-link exhibited considerably different masses with the different cross-linkers (BS^3 adds 138.07 Da per event, whereas DSP adds 173.98 Da, Figure 2c,d). When two peptides that each contain one Lys residue are connected by a cross-linker, it can be concluded that the ϵ-amino group of both Lys residues lie within 11.4 Å (the reachable space arm length of BS^3, the shorter of the two cross-linkers) in the native protein structure.

Optimization of ApoA-I Cross-Linking Conditions. Many studies have demonstrated a concentration-dependent oligomerization of apoA-I (*36*−*38*). It is generally believed that apoA-I self-associates at concentrations >0.1 mg/mL. To determine if concentrations below 0.1 mg/mL would yield pure monomeric forms of apoA-I for the analysis, we performed pilot cross-linking experiments at 0.1 mg/mL followed by ultrafiltrate concentration. However, 8−25% denaturing polyacrylamide gradient gel electrophoresis (SDS−PAGGE) and direct infusion mass spectroscopy measurements demonstrated that significant levels of dimeric and higher-order multimeric forms could still be detected in apoA-I cross-linked at 0.1 mg/mL (data not shown). Since there was no advantage in terms of studying the monomeric form in isolation and since working at this concentration required large volumes and nearly prohibitive amounts of expensive cross-linking reagent, we chose to cross-link apoA-I at 1 mg/mL and then separate the monomeric form from the higher-order oligomers by gel filtration after quenching the cross-linking reaction. This had the advantage of allowing us to isolate the dimeric form of the protein from the same reaction mixture as the monomer. Moreover, the protein concentration we used was compatible with other mass spectroscopic cross-linking protein structural studies carried out after higher-order aggregate separation (*39*).

An additional consideration was the possibility that the formation of initial cross-links may adversely affect the structure of the protein and potentiate the formation of additional cross-links that are not pertinent to the native protein conformation, seriously compromising any structural conclusions. Such an event could also be envisioned to promote nonspecific self-association of the protein. To address these concerns, we performed a series of cross-linking experiments with a range of apoA-I to cross-linker molar ratios from 1:5 to 1:50. Our reasoning was that at a low ratio of apoA-I to cross-linker, there would be fewer cross-links per apoA-I molecule compared to a high ratio. Therefore, the chance of cross-links to perturb the conforma-

[2] Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E.; Wang, J., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K. and Kollman, P. A. (2002) AMBER 7, University of California, San Francisco.

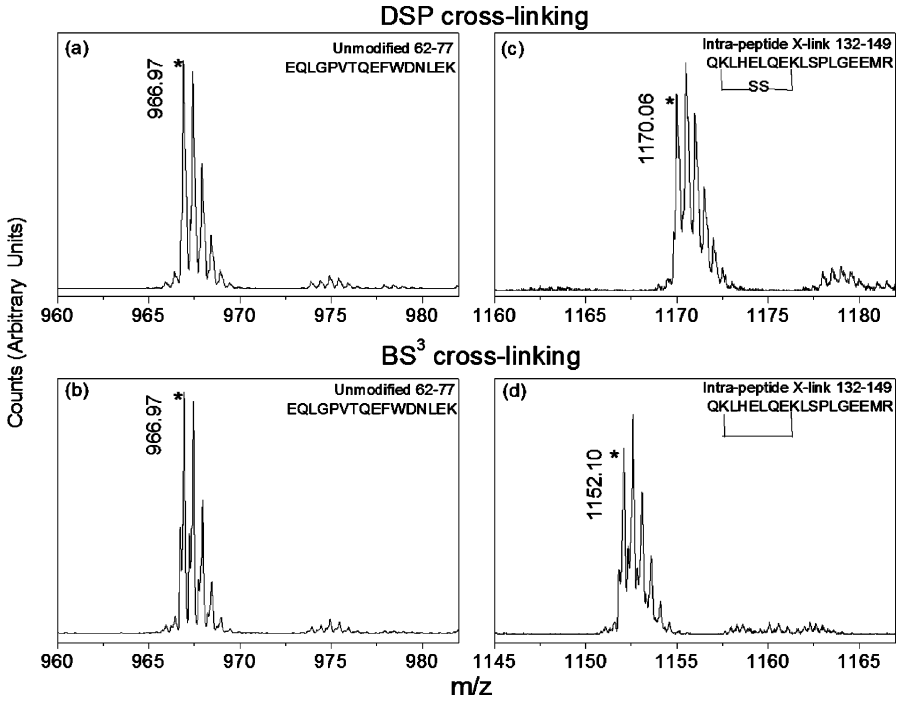

FIGURE 2: Effect of different cross-linkers of similar spacer arm length on the molecular weight of the resulting cross-linked peptides. Mass spectra of two peptides corresponding to 62−77 and 132−149 are shown from DSP (top) and BS^3 (bottom) cross-linking experiments. The monoisotopic peak of each mass cluster is indicated by an asterisk with its corresponding mass value. The peptide 62−77 exhibited identical mass values in both experiments (a) and (b) because it lacks a reactive Lys residue and was not modified by either cross-linker. On the other hand, the peptide 132−149, which contains two Lys residues, exhibited distinct masses (c) and (d) because of differing mass additions from the cross-linking agents. This mass difference is due to the use of two different cross-linkers of similar length, which eliminates the slight possibility of artifactual mass assignments as cross-linked peptides. (Note that only the doubly charged state for each peptide mass is shown.)

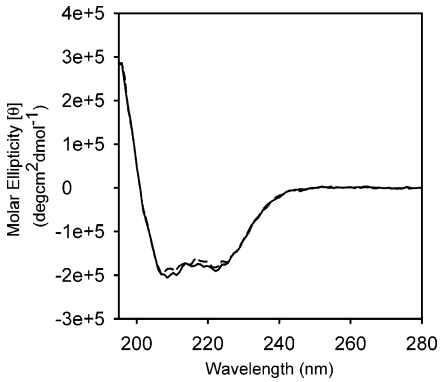

FIGURE 3: Far UV-circular dichroism spectra of unmodified (black) and cross-linked apoA-I (dotted). Cross-linking was carried out in PBS (pH 7.8) at 1 mg/mL protein concentration with 1:15 protein to cross-linker molar ratio. The conditions used are identical to mass spectroscopic sample preparations (see Table 1). The spectra were recorded in a 0.01-cm path cell with 1-s time constant averaging four scans at room temperature.

tion of the protein should be lower in the 1:5 sample vs the 1:50 sample. However, our MS experiments detected the same constellation of peptide cross-links in both low and high cross-linker to apoA-I ratios (data not shown). Furthermore, the band intensities corresponding to the higher-order self-associated forms of apoA-I remained the same in both ratios as detected by 8−25% SDS−PAGGE. This argues against the possibility of nonspecific protein self-association induced by the cross-linking. In addition, we measured the far UV−CD spectra of unmodified vs cross-linked apoA-I (Figure 3). The nearly identical spectra indicated that there are no significant secondary structural changes in cross-linked vs unmodified apoA-I recorded under the identical conditions as used for the mass spectroscopic studies. We also monitored the fluorescence emission of the four Trp residues present in apoA-I that was cross-linked under identical conditions as for the mass spectrometry experiments. At 4 °C, the unmodified apoA-I exhibited a wavelength of maximum fluorescence (λ-max) of 333.2 ± 0.6 nm. The cross-linked protein exhibited a λ-max of 333.3 ± 0.3 nm, indicating that the cross-linking reaction did not perturb the environment of the Trp residues. There was also no difference at 25 °C. Taken together, we saw no evidence that the cross-linker induced spurious changes in protein conformation, even at higher cross-linker concentrations.

Mass Spectrometry. The data presented in this paper are based on three different cross-linking experiments in which either the cross-linker type or molar ratio was varied (apoA-I:BS3 1:5, apoA-I:BS3 1:15, and apoA-I:DSP 1:5). The different stages of a typical cross-linking experiment are shown by the 8−25% SDS−PAGGE analysis shown in Figure 4. It is clear that apoA-I cross-linked in a range of oligomers that were roughly evenly spread between monomer, dimer, trimer, and tetramer with higher-order oligomers present (lane 2, Figure 4). We were able to effectively separate the monomeric and dimeric forms by gel filtration (lanes 3 and 4, Figure 4).

Tables 2 and 3 summarize the identified peptide cross-links within the monomeric form of cross-linked apoA-I. The peptides in Table 2 are restricted to 10 "intrapeptide" cross-links and are defined to occur between two Lys residues located on the same tryptic peptide. Although these are of relatively short range and provide little information on the

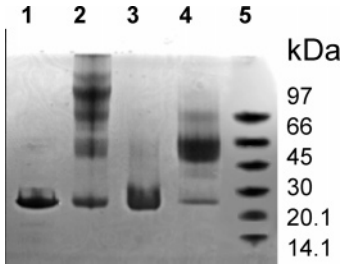

FIGURE 4: Cross-linked apoA-I and chromatographically separated cross-linked apoA-I components analyzed on a 8−25% denaturating SDS Phast gel. Unmodified apoA-I (lane 1), BS^3 cross-linked apoA-I (lane 2), cross-linked monomeric and dimeric apoA-I after gel filtration chromatographic separation (lanes 3 and 4, respectively), the low molecular weight standard (lanes 5). The gel was stained with Coomassie Blue.

3-D structure of the protein, because the cross-linking of Lys residues is dependent on their spatial arrangement in addition to their distance apart within the protein sequence, even cross-links that span 2−3 amino acids can provide useful information on protein backbone arrangement.

Table 3 lists the 7 "interpeptide" cross-links that we identified. These are defined as peptides within the same molecule of apoA-I, but are present on different tryptic peptides, i.e., they contain at least one trypsin cleavage site between the peptides. (Note: The term "interpeptide" should not be confused with "intermolecular" cross-links, which refer to cross-linked Lys residues located on two separate molecules of apoA-I in the case of the cross-linked dimer sample, see below). These interpeptide cross-links tend to be longer in range with respect to the amino acid sequence and offer the most information in terms of 3-D structure. The set of cross-links shown in Tables 2 and 3 were identified in each of the three experiments performed. The fact that the BS^3 and DSP experiments identified the same cross-links indicates that the small amount of DMSO used in the DSP cross-linking did not affect apoA-I conformation. The excellent agreement between the theoretical and the experimental mass values indicates the high accuracy of experimental mass determination. Not surprisingly, in most cases, masses bearing intrapeptide cross-links were more abundant than interpeptide cross-links consistent with the higher probability of their formation. It was apparent that some Lys residues were more prone to cross-linker reactions than others. For example, of all 21 Lys residues in apoA-I, only K182 was not detected in any kind of cross-linker modification, suggesting that this residue may be buried within the other amino acid side chains. The only other comparatively inert lysine residue was K77, from which only a fraction was detected as modified with a hydrolyzed cross-link. This may indicate that K77 is accessible to the cross-linker, but another lysine may not be within the reachable spacer arm length for complete cross-linker formation. All the remaining 19 Lys residues either contained hydrolyzed cross-linkers and/or participated in successful cross-links. K96 was the most active being involved in one intra- and three different interpeptide cross-links. This is in marked contrast to the our previous studies of the lipid-bound form of apoA-I from in which the most active sites identified were K118, K133, and K140 (*12*).

Homology Model. A schematic representation illustrating the locations of the cross-links in Tables 2 and 3 is shown in Figure 5. This information was used to constrain the building of an all atom, 3-D model of lipid-free monomeric apoA-I using sequence homology and energy minimization techniques. The homology model was based on four templates chosen from the PDB to model different regions of apoA-I are highlighted with different colors in Figure 5. These partial models were then assembled into a combined model subject to all 17 cross-link constraints derived from this study. The spacer arm length was restricted to 11.4 Å, the shorter of the two cross-linkers used. From the available templates with few sequence alignment gaps, we were able to obtain reliable predictions for the regions 15−187 and 200−241 of apoA-I. However, the regions from 1 to 14 and 188−199 lacked a suitable template and were arbitrarily modeled as random coil, but it remains possible that these regions may contain some secondary structure. The predicted α-helical regions for apoA-I are shown in Figure 6, compared to the positions of repeating amphipathic helices that were predicted from sequence analysis.

As seen in the ribbon representation of the model (Figure 7A) the protein can generally be viewed as a bundle of at least four helices. The first is a relatively short helix between residues 24−39. The second follows after a reverse turn and contains 48−82 with a short break centered at ∼L64. After a loop region, the third helix encompasses residues 100−129 again with a short discontinuity at ∼P120 and turns near K133 into the antiparallel helix 4 encompassing residues 134−164. In addition, there are two short helices in the C-terminal third of the molecule. Overall, the protein is well packed as viewed by the space-filling illustration (Figure 7B), indicating maintenance of hydrophobic interactions and a hydrophobic core. However, there is some exposure of helical regions to the solvent, which could account for the lower stability of lipid-free apoA-I compared to other globular proteins (*15*, *40*). Finally, a key prediction that resulted from the modeling was that the N- and C-termini of apoA-I are situated in relatively close proximity.

We also investigated the cross-linking pattern in the isolated dimeric form of lipid-free apoA-I. Six intermolecular cross-links that were unique to the dimer mass spectra were identified by comparison with the monomer spectra (Table 4) (*12*). All the cross-links available in the monomer (Tables 2 and 3) were found to be present in the dimer sample except two, which are also listed in Table 4. Interestingly, the majority of the cross-links that were unique to the dimer spectra involved either the N- or C-terminal regions of apoA-I.

DISCUSSION

In the current study, we have generated the most comprehensive list of distance constraints for the monomeric form of lipid-free apoA-I to date. A major strength of our approach is the inherent random nature of the cross-linking reaction among numerous Lys residues in apoA-I. This has two advantages over widely used energy transfer strategies that have been used previously. First, our approach did not require the introduction of potentially disruptive mutations to add large probes to the molecule. This allowed the direct study of the native human plasma form of apoA-I. Second, the random nature of the cross-linking reaction allowed us to measure 17 different distance constraints simultaneously within the native protein, whereas energy transfer experi-

Table 2: Intrapeptide Cross-Links Identified in Internally Cross-Linked apoA-I Monomer

X-linked peptide residues[a]	peptide involved	mass observed[b] (Da)	mass calculated (Da)	deviation[c] (Da)	distance[d] (Å)
K12−K23	11−27	2015.10	2015.10	0.00	11.0
K40−K45*	28−59	3728.02	3727.86	0.16	9.5
K88−K94	84−96	1671.84	1671.84	0.00	10.4
K94−K96*[e]	89−106	2302.19	2302.18	0.01	10.4
K96−K106	95−107	1716.93	1716.92	0.01	11.4
K106−K107*	97−116	2782.44	2782.44	0.00	6.5
K133−K140	132−149	2302.19	2302.20	0.01	11.4
K206−K208*	196−215	2346.25	2346.25	0.00	3.8
K226−K238	216−239	2863.63	2863.63	0.00	11.4
K238−K239*	227−243	2108.12	2108.11	0.01	6.9

[a] Cross-linked peptides shown by asterisks can be considered short range with respect to the protein sequence (see Results). [b] Observed monoisotopic masses were obtained from apoA-I/BS3 1:5 cross-linking experiments. However, all cross-links were observed in all three experiments (see Results). [c] Deviation = observed mass − expected mass [d] The measured distance between ϵ-amino groups of two lysines involved in cross-linking; taken from homology model (Figure 7). [e] Presence of both cross-linked peptides K94−K96 and K133−K140 with a similar monoisotopic mass 2302.19 Da was confirmed by their different retention times 51.8 and 56.7 min during RP-HPLC separation.

Table 3: Interpeptide Cross-Links Identified in Internally Cross-Linked apoA-I Monomer

X-linked peptide residues	peptides involved	mass observed[a] (Da)	mass calculated (Da)	deviation[b] (Da)	distance[c] (Å)
N_T-K96[d]	1−10 95−106	2814.47	2814.35	0.12	11.4
K23−K59	13−27 46−61	3669.01	3668.90	0.11	7.5
K94−K208	89−96 207−215	2080.07	2080.14	0.07	4.5
K96−K195	95−106 189−206	3616.02	3615.85	0.16	10.3
K96−K208[e]	95−106 207−215	2630.30	2630.30	0.00	7.2
K96−K226	95−106 216−238	4186.28	4186.21	0.07	11.4
K118−K140	117−123 134−149	2914.59	2914.56	0.03	11.4

[a] Observed monoisotopic masses were obtained from the apoA-I/BS3 1:5 cross-linking experiment. However, all cross-links were observed in all three experiments (see Results). [b] Absolute difference between the observed and the calculated monoisotopic masses [c] The measured distance between ϵ-amino groups of two lysines involved in cross-linking; taken from homology model (Figure 7). [d] The term N_T represents the extreme N-terminal amino group. [e] The experimental mass corresponding to the cross-link K96−K208 was taken from the apoA-I/DSP 1:5 cross-linking experiment due to a technical anomaly that precluded obtaining an accurate mass from the peak in apoA-I/BS3 1:5 cross-linking experiment.

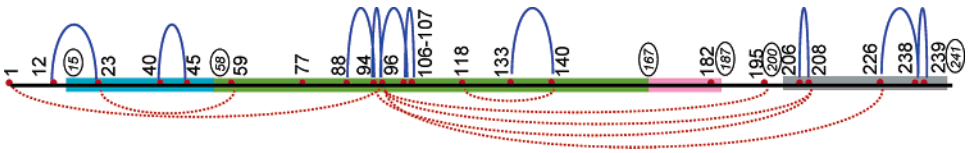

FIGURE 5: Schematic representation of intrapeptide and interpeptide cross-links identified in cross-linked monomeric apoA-I. The protein is presented as a black solid line on which each lysine residue and N-terminus amino group are shown as filled circles with numbers corresponding to their location in the protein sequence. The four protein templates used for modeling are shown in different colors (Ng Gtpase domain, blue; apolipoprotein E4, green; apoA-I, pink; enoyl acyl carrier protein reductase, gray) overlaid with the black solid line. Boundaries of the regions of apoA-I represented by each template is marked by two circled italicized numbers (see Table 1 for more details). The location of intrapeptide and interpeptide cross-links (listed in Tables 2 and 3) are shown as solid and dotted curved lines, respectively.

ments require a different set of mutants for each individual distance being measured. Using our cross-link map in conjunction with structural homology modeling, we built the first all atom molecular model of lipid-free apoA-I available to date. Below, we compare the model's predictions to currently available experimental structural evidence for lipid-free apoA-I and discuss the strengths and limitations of the model.

The Model vs Experimental Data. The model put forward (Figure 7) was based solely on the cross-link distance constraints, homology modeling based on amphipathic helicity and energy minimization calculations. No other experimental data was factored into the production of the model. Therefore, the most robust test of the validity of the basic model lies in how well its predictions fit experimental evidence contained in the literature. An obvious starting point is to compare the predicted total secondary structure content of the model with experimental secondary structure estimations measured by circular dichroism spectroscopy. The overall α-helicity of the model is ∼56% (136/243 residues) as assigned by the MOE. We performed a secondary structure content analysis on lipid-free apoA-I by far UV−CD as described in Experimental Procedures. The analysis provided an excellent theoretical fit to the experimental data and estimated a helical content of 55% α-helix, 8% β-sheet, and 37% coil/turns. Therefore, the helical content predicted by the model matches quite well with the experimental value. Sequence analysis suggests that short runs of β-sheet may

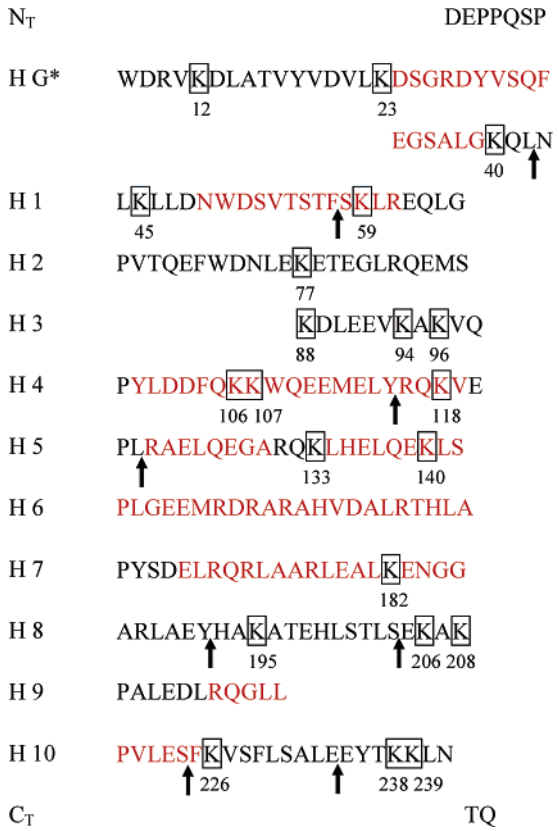

FIGURE 6: Secondary structure content of the homology model. The sequence is arranged according to the putative amphipathic helical segments that were predicted by sequence analysis are indicated by H1 through H10 and N-terminal 43 residues and C-terminal two residues are shown separately (from Roberts et al. ref *21*). The predicted α-helical regions from the current sequence threading model are shown in red. The proteolytically sensitive sites observed by previous limited proteolytic digestion experiments performed on lipid-free apoA-I (*21*, *22*) are indicated by arrows.

exist in apoA-I, especially in the N-terminus (*41*). Although we were unable to model residues 1−14 due to lack of a homology template, it remains possible that this region may contain *β*-sheet. Even if there is some additional helical content within these domains, the overall helicity would still not be far out of line with the CD secondary structure estimates.

The model also agrees with structural conclusions drawn from limited proteolysis studies. Roberts et al. (*21*) and Ji et al. (*22*) identified several proteolytically sensitive sites within lipid-free apoA-I. The assumption is that these sites represent highly solvated, mobile regions that are not part of organized secondary structural elements. These sites are indicated by arrows in Figure 6. Interestingly, each site exists in a region assigned as a random structure (a loop or a turn) or close to ends of a predicted helical region in our model. Only one of these sites (Y191) was located in a region that was arbitrarily assigned to random structure due to lack of a homology template. All the rest coincide with random coil as a natural result of the distance constraints created by cross-linking. In addition, the model agrees well with previously measured distance constraints determined by resonance energy transfer methods. For example, Oda et al. used electron paramagnetic resonance spin coupling (EPR) to determine that probes attached to residues L163, G217, and K226 are proximal within the protein (within 22 Å) of each other in lipid-free apoA-I (*42*). In our model, these residues (considering the nearest heavy atoms) are all situated within a radius of 20 Å. Furthermore, Agree et al. (*18*) used fluorescence energy transfer to demonstrate that probes attached to the C-terminal residues 190 and 232 are in close proximity to the four tryptophan donors located in the N-terminal half of the protein. This has also been suggested by other studies (*19*, *40*). This is consistent with the proximity of the N- and C-termini apparent in Figure 7.

The model is also consistent with the degree of availability of Lys residues to the cross-linking agents. For example, the most unreactive lysine K182 does not have any other nearby lysines for cross-linking and is somewhat buried within molecule in the model. On the other hand K77, the other unreactive lysine, although it does have a single potential cross-link target within range (K88), it is modeled in an unfavorable geometry to complete the cross-link. In contrast to K182, K77 is on the surface of the molecule, possibly explaining why we observed a hydrolyzed cross-linker at K77 and not at K182.

The model also suggests predictions that are somewhat surprising. Since the sequence of apoA-I was reported, it has been speculated that the regularly spaced proline residues may act as helix breakers and promote the formation of turns in the structure. This proved not to be likely in the lipid-bound form of apoA-I (*43*) but may be important in the lipid-free form. However, we found little evidence that the turns within the lipid-free model are centered on proline residues. More often than not, prolines are modeled within a helix kink or random coil as opposed to turn sequences. This suggests that, whatever the function of the proline residues, it is apparently not to induce helical turns. Furthermore, we found only partial agreement of cross-linker distance constraints with a Δ(1−43) monomer unit extracted from the tetramer molecular assembly from the apo-I Δ(1−43) crystal structure (*14*). This suggests that the interactions that stabilize the tetrameric form of the protein under the conditions of the X-ray crystallography studies may be different from those occurring in the monomeric form under native buffer conditions.

In addition to the monomeric form of the protein, we also gained insight into the nature of the self-association of apoA-I in solution from the analysis of the dimeric form. Although we have not yet attempted detailed modeling of the dimer, two findings are worthy of note. First, we found that almost all of the cross-links identified within the monomer were also present in the dimer (Table 4). This suggests that the protein associates with the majority of the monomeric interactions intact without a significant unfolding or conformational change. Second, with the exception of the disappearance of the cross-link at K118−K140 from the dimeric spectra, most of the differences between the cross-links in the monomer and dimer spectra were localized to about the N- and C-terminal 50 amino acids of the protein. Therefore, we propose that apoA-I may dimerize by allowing the random coil regions in both the N- and C-termini molecule to interact with the C- and N-regions, respectively, in the second molecule as shown in Figure 8. As we have suggested before (*44*), this interaction may allow for the stabilization of amphipathic helical domains that are unfolded in the absence of the second molecule of apoA-I. It is worthwhile to note that a potentially similar C-to-N interaction has been documented in lipidated HDL particles (*12*).

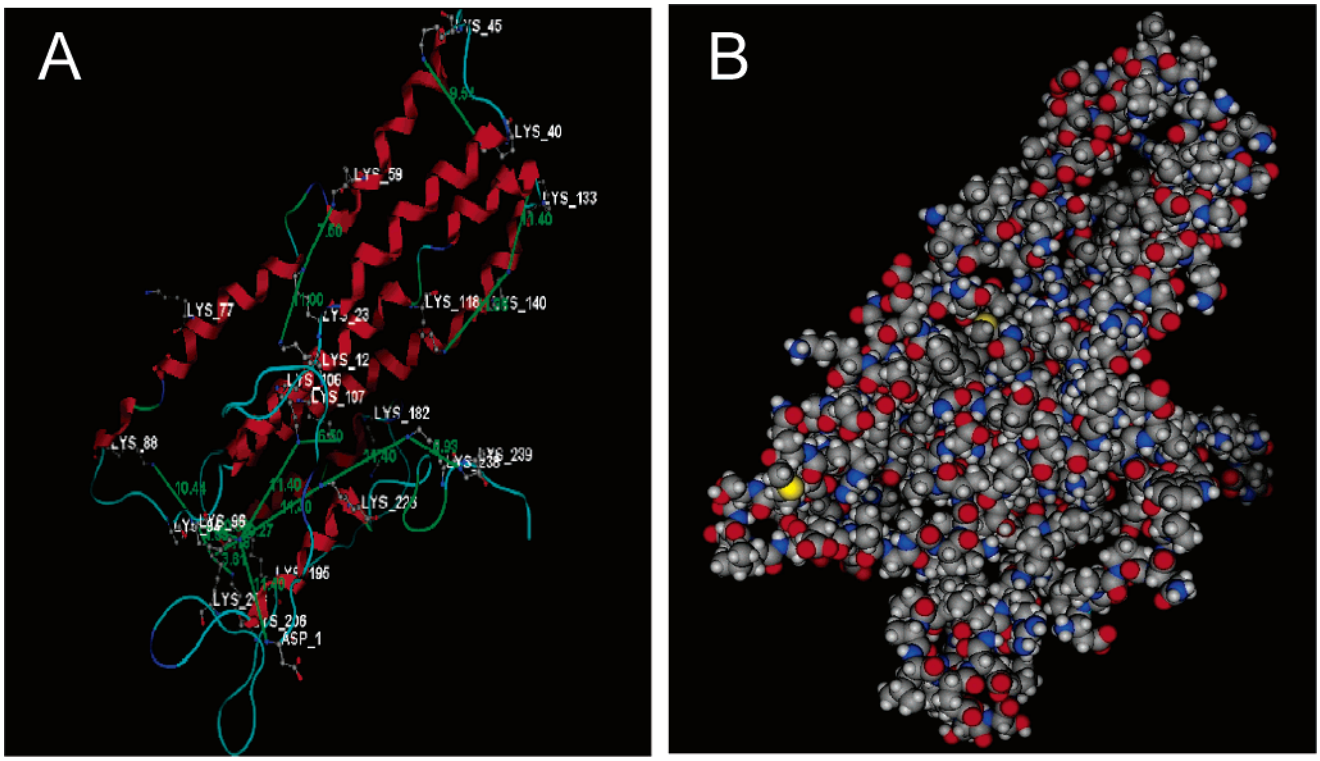

FIGURE 7: Ribbon and space filling model representations of the homology model created for lipid free apoA-I based on cross-link distance constraints. In the ribbon diagram (A) secondary structural types are color-coded as α-helix (red), random coil (light blue), and turns (dark blue and green). Each lysine residue is shown (stick representation) with its corresponding position within the protein sequence. Observed cross-links are presented in green with their distances in angstroms. These distances are also listed in the last column of Tables 2 and 3. In the space filling model (B), the atoms are color coded as follows: carbon and hydrogen (grey) oxygen (red), nitrogen (blue), and sulfur (yellow).

Table 4: Intermolecular Cross-Links Pertinent to the apoA-I Cross-Linked Dimer

X-linked peptide residues	peptides involved	mass observed[a] (Da)	mass calculated (Da)	deviation[b] (Da)
	Cross-Links Unique to the apoA-I Dimer Spectra			
N_T-N_T[c]	1−10 1−10	2589.29	2589.14	0.15
N_T-K77	1−10 62−83	3980.91	3980.87	0.04
N_T-K238	1−10 227−239	2877.51	2877.41	0.10
K208−K208	207−215 207−215	2161.20	2161.21	0.01
K208−K238	207−215 227−239	2663.59	2663.44	0.15
K226/K238−K238	216−238 227−239	4249.35	4249.27	0.08
	Cross-Links Unique to the apoA-I Monomer Spectra			
K226−K238	216−239		2863.63	
K118−K140	117−123 134−149		2914.56	

[a] Observed monoisotopic masses were obtained from apoA-I/BS3 1:5 cross-linking experiment. However, all cross-links were observed in all three experiments (see Results). [b] Deviation = observed mass − expected mass. [c] The term N_T represents the extreme N-terminal amino group.

Although our current experiments did not address the higher oligomers, it is conceivable that these could form by a staggered arrangement with the same N-to-C interactions.

Limitations of the Model. There is no question that apoA-I is a thermodynamically unstable and highly dynamic molecule when present in the lipid-free form. Hence, any detailed model must be interpreted with a healthy respect for conformational motions undoubtedly occurring within the structure. The cross-linking approach that we have used in this study is limited for studying dynamic systems because the results of the experiment require the use of all the measured distance constraints in only a single model. For example, if two conformers of apoA-I are present in solution as has been suggested (*20*), this methodology would not be able to distinguish between the two. The result would be a set of cross-links that reflect all the structures in a system, not unlike most spectroscopic techniques. The strength of the method, however, lies in the ability to identify regions in the protein that are in close proximity for a long enough period of time to form the cross-link. Therefore, the method should result in a basic model that represents the most stable or longest lived conformation of the protein. In terms of

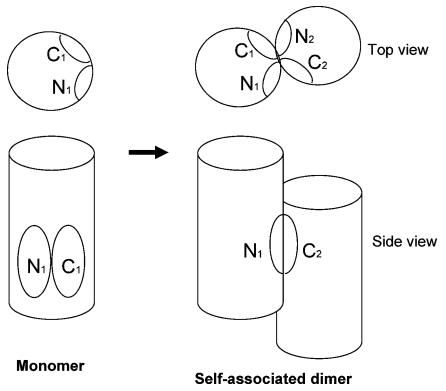

FIGURE 8: A basic model representation for dimeric self-association of apoA-I. The compact apoA-I monomer is presented abstractly as a cylinder with N- and C-termini together as suggested by our all atom model (Figure 7). The association of the N- and C-termini are shown interacting in the dimer with most of the helical bundle structure of each apoA-I constituent remaining intact.

overall geometry, our model (70 × 35 Å) agrees generally with the compact form of lipid-free apoA-I observed by Rogers et al. (*20*) with dimensions of about 100 × 28 Å, based on their sedimentation velocity experiments. However, our data does not preclude the existence of an elongated form, especially if it is a relatively short-lived intermediate. The model should be considered an averaged snapshot of a conformation that the protein can adopt in solution.

It is also important to note that typical homology modeling is performed using a single template molecule that is reasonably related to the protein being modeled with few gaps in the alignment. Because of the relative paucity of high-definition structural information for apolipoproteins, we were forced to build model in segments using different templates for particular runs of sequence. Inevitably, there were regions for which we had no usable template and had to essentially fill in the gaps with random coil. Thus, we may have missed some subtleties in the structure that may prove to be functionally important. Fortunately, the cross-linking constraints provided sufficient information to allow the reliable placement of the individual segments within the combined model.

Despite these limitations, the model is the first to be constructed for the entire sequence of apoA-I using 3-D distance constraints as an integral component of model building. Although the resolution is significantly lower than X-ray or NMR-derived structures, the model provides a sufficient level of detail for the design of future resonance energy transfer experiments that can confirm the predicted spatial relationships of particular regions of the protein. Indeed, experiments are currently underway to test the predicted proximity of the N- and C-termini. Most importantly, with the advancements in our understanding of lipid-bound apoA-I conformation, this model will form a basis for deciphering the conformational transitions of apoA-I as it binds lipids.

ACKNOWLEDGMENT

We thank Dr. Sissel Lund-Katz from the Children's Hospital of Philadelphia for providing the highly calibrated circular dichroism scans for lipid-free apoA-I.

REFERENCES

1. *1997 Heart and Stroke Statistical Update. 12/96*, American Heart Association, Dallas, 1997.
2. Jonas, A. (1998) Regulation of lecithin cholesterol acyltransferase activity, *Prog. Lipid Res. 37*, 209−234.
3. Xu, S., Laccotripe, M., Huang, X., Rigotti, A., Zannis, V. I., and Krieger, M. (1997) Apolipoproteins of HDL can directly mediate binding to the scavenger receptor SR−BI, an HDL receptor that mediates selective lipid uptake, *J. Lipid Res. 38*, 1289−1298.
4. Oram, J. F. (2003) HDL apolipoproteins and ABCA1: partners in the removal of excess cellular cholesterol, *Arterioscler. Thromb. Vasc. Biol. 23*, 720−727.
5. Brouillette, C. G., Anantharamaiah, G. M. (1995) Structural models of human apolipoprotein A−I, *Biochim. Biophys. Acta 1256*, 103−129.
6. Klon, A. E., Segrest, J. P., and Harvey, S. C. (2002) Molecular dynamics simulations on discoidal HDL particles suggest a mechanism for rotation in the apo A−I belt model, *J. Mol. Biol. 324*, 703−721.
7. Klon, A. E., Segrest, J. P., and Harvey, S. C. (2002) Comparative models for human apolipoprotein A-I bound to lipid in discoidal high-density lipoprotein particles, *Biochemistry 41*, 10895−10905.
8. Segrest, J. P., Jones, M. K., Klon, A. E., Sheldahl, C. J., Hellinger, M., De Loof, H., and Harvey, S. C. (1999) A detailed molecular belt model for apolipoprotein A-I in discoidal high-density lipoprotein, *J. Biol. Chem. 274*, 31755−31758.
9. Koppaka, V., Silvestro, L., Engler, J. A., Brouillette, C. G., and Axelsen, P. H. (1999) The structure of human lipoprotein A-I. Evidence for the "belt" model, *J. Biol. Chem. 274*, 14541−14544.
10. Maiorano, J. N., and Davidson, W. S. (2000) The orientation of helix 4 in apolipoprotein A-I-containing reconstituted high-density lipoproteins, *J. Biol. Chem. 275*, 17374−17380l.
11. Panagotopulos, S. E., Horace, E. M., Maiorano, J. N., and Davidson, W. S. (2001) Apoliprotein A-I adopts a belt-like orientation in reconstituted high-density lipoproteins, *J. Biol. Chem. 277*, 39477−39484.
12. Davidson, W. S., and Hilliard, G. M. (2003) The spatial organization of apolipoprotein A-I on the edge of discoidal high-density lipoprotein particles: A mass spectrometry study, *J. Biol. Chem. 278*, 27199−27207.
13. Li, H., Lyles, D. S., Thomas, M. J., Pan, W., and Sorci-Thomas, M. G. (2000) Structural determination of lipid-bound ApoA-I using fluorescence resonance energy transfer, *J. Biol. Chem. 275*, 37048−37054.
14. Borhani, D. W., Rogers, D. P., Engler, J. A., and Brouillette, C. G. (1997) Crystal structure of truncated human apolipoprotein A-I suggests a lipid-bound conformation, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12291−12296.
15. Fang, Y., Gursky, O., and Atkinson, D. (2003) Structural studies of N- and C-terminally truncated human apolipoprotein A-I, *Biochemistry 42*, 6881−6890.
16. Rogers, D. P., Brouillette, C. G., Engler, J. A., Tendian, S. W., Roberts, L., Mishra, V. K., Anantharamaiah, G. M., Lund-Katz, S., Phillips, M. C., and Ray, M. J. (1997) Truncation of the amino terminus of human apolipoprotein A-I substantially alters only the lipid-free conformation, *Biochemistry 36*, 288−300.
17. Wang, G., Sparrow, J. T., and Cushley, R. J. (1997) The helix-hinge-helix structural motif in human apolipoprotein A-I determined by NMR spectroscopy, *Biochemistry 36*, 13657−13666.
18. Behling Agree, A. K., Tricerri, M. A., Arnvig, M. K., Tian, S. M., and Jonas, A. (2002) Folding and stability of the C-terminal half of apolipoprotein A-I examined with a Cys-specific fluorescence probe, *Biochim. Biophys. Acta 1594*, 286−296.
19. Tricerri, M. A., Behling Agree, A. K., Sanchez, S. A., and Jonas, A. (2000) Characterization of apolipoprotein A-I structure using a cysteine- specific fluorescence probe, *Biochemistry 39*, 14682−14691.
20. Rogers, D. P., Roberts, L. M., Lebowitz, J., Engler, J. A., and Brouillette, C. G. (1998) Structural analysis of apolipoprotein A-I: effects of amino- and carboxy-terminal deletions on the lipid-free structure, *Biochemistry 37*, 945−955.
21. Roberts, L. M., Ray, M. J., Shih, T. W., Hayden, E., Reader, M. M., and Brouillette, C. G. (1997) Structural analysis of apolipoprotein A−I: limited proteolysis of methionine-reduced and -oxidized lipid-free and lipid-bound human apo A-I, *Biochemistry 36*, 7615−7624.

22. Ji, Y., and Jonas, A. (1995) Properties of an N-terminal proteolytic fragment of apolipoprotein AI in solution and in reconstituted high-density lipoproteins, *J. Biol. Chem. 270*, 11290−11297.
23. Calabresi, L., Meng, Q. H., Castro, G. R., and Marcel, Y. L. (1993) Apolipoprotein A-I conformation in discoidal particles: evidence for alternate structures, *Biochemistry 32*, 6477−6484.
24. Gursky, O., and Atkinson, D. (1996) Thermal unfolding of human high-density apolipoprotein A-1: implications for a lipid-free molten globular state, *Proc. Natl. Acad. Sci. U.S.A. 93*, 2991−2995.
25. Gorshkova, I. N., Liu, T., Zannis, V. I., and Atkinson, D. (2002) Lipid-free structure and stability of apolipoprotein A-I: probing the central region by mutation, *Biochemistry 41*, 10529−10539.
26. Saito, H., Dhanasekaran, P., Nguyen, D., Holvoet, P., Lund-Katz, S., and Phillips, M. C. (2003) Domain structure and lipid interaction in human apolipoproteins A-I and E, a general model, *J. Biol. Chem. 278*, 23227−23232.
27. Wilson, C., Wardell, M. R., Weisgraber, K. H., Mahley, R. W., and Agard, D. A. (1991) Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E, *Science 252*, 1817−1822.
28. Wang, J., Sykes, B. D., and Ryan, R. O. (2002) Structural basis for the conformational adaptability of apolipophorin III, a helix-bundle exchangeable apolipoprotein, *Proc. Natl. Acad. Sci. U.S.A. 99*, 1188−1193.
29. Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., and Dollinger, G. (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry, *Proc. Natl. Acad. Sci. U.S.A. 97*, 5802−5806.
30. Rahaman, A., Srinivasan, N., Shamala, N., and Shaila, M. S. (2003) The fusion core complex of the peste des petits ruminants virus is a six-helix bundle assembly, *Biochemistry 42*, 922−931.
31. Jaspard, B., Collet, X., Barbaras, R., Manent, J., Vieu, C., Parinaud, J., Chap, H., and Perret, B. (1996) Biochemical characterization of pre-beta 1 high-density lipoprotein from human ovarian follicular fluid: evidence for the presence of a lipid core, *Biochemistry 35*, 1352−1357.
32. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926−935.
33. Compton, L. A., and Johnson, W. C., Jr. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication, *Anal. Biochem. 155*, 155−167.
34. Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem. 287*, 252−260.
35. Bennett, K. L., Kussmann, M., Bjork, P., Godzwon, M., Mikkelsen, M., Sorensen, P., and Roepstorff, P. (2000) Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping—a novel approach to assess intermolecular protein contacts. *Protein Sci. 9*, 1503−1518.
36. Barbeau, D. L., Jonas, A., Teng, T., and Scanu, A. M. (1979) Asymmetry of apolipoprotein A-I in solution as assessed from ultracentrifugal, viscometric, and fluorescence polarization studies, *Biochemistry 18*, 362−369.
37. Vitello, L. B., and Scanu, A. M. (1976) Studies on human serum high-density lipoproteins. Self-association of apolipoprotein A-I in aqueous solutions, *J. Biol. Chem. 251*, 1131−1136.
38. McGuire, K. A., Davidson, W. S., and Jonas, A. (1996) High yield overexpression and characterization of human recombinant proapolipoprotein A-I, *J. Lipid Res. 37*, 1519−1528.
39. Kruppa, G. H., Schoeniger, J., and Young, M. M. (2003) A top down approach to protein structural studies using chemical cross-linking and Fourier transform mass spectrometry, *Rapid Commun. Mass Spectrom. 17*, 155−162.
40. Gorshkova, I. N., Liadaki, K., Gursky, O., Atkinson, D., and Zannis, V. I. (2000) Probing the lipid-free structure and stability of apolipoprotein A-I by mutation, *Biochemistry 39*, 15910−15919.
41. Nolte, R. T., and Atkinson, D. (1992) Conformational analysis of apolipoprotein A-I and E-3 based on primary sequence and circular dichroism, *Biophys. J. 63*, 1221−1239.
42. Oda, M. N., Forte, T. M., Ryan, R. O., and Voss, J. C. (2003) The C-terminal domain of apolipoprotein A-I contains a lipid-sensitive conformational trigger, *Nat. Struct. Biol. 10*, 455−460.
43. Brouillette, C. G., Anantharamaiah, G. M., Engler, J. A., and Borhani, D. W. (2001) Structural models of human apolipoprotein A-I: a critical analysis and review, *Biochim. Biophys. Acta 1531*, 4−46.
44. Davidson, W. S., Hazlett, T., Mantulin, W. W., and Jonas, A. (1996) The role of apolipoprotein AI domains in lipid binding, *Proc. Natl. Acad. Sci. U.S.A. 93*, 13605−13610.

BI047717+