THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 29, Issue of July 18, pp. 27199–27207, 2003

Printed in U.S.A.

The Spatial Organization of Apolipoprotein A-I on the Edge of Discoidal High Density Lipoprotein Particles

A MASS SPECTROMETRY STUDY*

Received for publication, March 18, 2003, and in revised form, April 28, 2003
Published, JBC Papers in Press, April 30, 2003, DOI 10.1074/jbc.M302764200

W. Sean Davidson‡§ and George M. Hilliard¶

From the ‡Department of Pathology and Laboratory Medicine, University of Cincinnati, Cincinnati, Ohio 45267-0529 and the ¶Department of Molecular Sciences and Center of Excellence in Genomics and Bioinformatics, University of Tennessee Health Science Center, Memphis, Tennessee 38163

The three-dimensional structure of human apoA-I on nascent, discoidal HDL particles has been debated extensively over the past 25 years. Recent evidence has demonstrated that the α-helical domains of apoA-I are arranged in a belt-like orientation with the long axis of the helices perpendicular to the phospholipid acyl chains on the disc edge. However, experimental information on the spatial relationships between apoA-I molecules on the disc is lacking. To address this issue, we have taken advantage of recent advances in mass spectrometry technology combined with cleavable cross-linking chemistry to derive a set of distance constraints suitable for testing apoA-I structural models. We generated highly homogeneous, reconstituted HDL particles containing two molecules of apoA-I. These were treated with a thiol-cleavable cross-linking agent, which covalently joined Lys residues in close proximity within or between molecules of apoA-I in the disc. The cross-linked discs were then exhaustively trypsinized to generate a discrete population of peptides. The resulting peptides were analyzed by liquid chromatography/mass spectrometry before and after cleavage of the cross-links, and resulting peaks were identified based on the theoretical tryptic cleavage of apoA-I. We identified at least 8 intramolecular and 7 intermolecular cross-links in the particle. The distance constraints are used to analyze three current models of apoA-I structure. The results strongly support the presence of the salt-bridge interactions that were predicted to occur in the "double belt" model of apoA-I, but a helical hairpin model containing the same salt-bridge docking interface is also consistent with the data.

High plasma levels of high density lipoprotein (HDL)[1] are widely thought to be protective against human cardiovascular disease. Apolipoprotein (apo)A-I, a 243-amino acid, 28-kDa protein is a key mediator of HDL function. It is required for lecithin:cholesterol acyl transferase-mediated maturation of HDL and may be a major ligand by which cholesteryl esters are delivered to the liver via the scavenger receptor type B class 1 receptor (1). In addition, the recent discovery of the importance of the ATP binding cassette protein in lipid transport (2–4) indicates that an important apoA-I to cell surface interaction may occur during HDL formation and reverse cholesterol transport. One of the major obstacles to a better understanding of these interactions has been the paucity of detailed structural information for apoA-I in its various states of lipid association.

Homogeneous discoidal forms of HDL are easily reconstituted from purified protein and lipids *in vitro* (5), and these reconstituted HDL (rHDL) analogs have been used heavily for structural studies. The discs likely exist as a phospholipid/cholesterol bilayer surrounded at its edges by the hydrophobic regions of the amphipathic helices of apoA-I (for a review see Ref. 6). Segrest *et al.* and others (7, 8) proposed in the late 1970s that the α-helices of apoA-I were arranged around the disc circumference with the long axis of the helices perpendicular to the acyl chains. This became known as the "belt" or "bicycle wheel" model. Alternatively, other investigators theorized that the 22-amino acid helical repeats could traverse the bilayer edge with the helices parallel to the acyl chains (9). This "picket fence" model was challenged by the first successful x-ray crystal structure of a lipid-free fragment of apoA-I by Borhani *et al.* (10). The crystal structure showed a tetramer of highly α-helical apoA-I molecules arranged in a ring-shaped complex, with no evidence of hairpin turns. Borhani *et al.* hypothesized that the ring motif in the crystal structure could be applied to the case of lipid-bound apoA-I on a disc. Since then, a belt-like orientation for the helices of apoA-I has been supported by polarized infrared spectroscopy experiments performed by Koppaka *et al.* (11). In addition, we published a series of studies in which various single tryptophan mutants of apoA-I were analyzed in discoidal HDL particles containing phospholipids with quenching groups at various positions along the acyl chain. The results clearly showed that all eight 22-amino acid helices in apoA-I were oriented perpendicular to the phospholipid acyl chains (12, 13).

With the question of apoA-I helical orientation addressed, attention has focused on determining spatial relationships between two molecules of apoA-I on a disc. Segrest *et al.* (14) recently published a computer model referred to as the "double

* This work was supported in part by RO1 Grants HL62542 and HL67093 from NHLBI, National Institutes of Health (to W. S. D.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ An Established Investigator of the American Heart Association. To whom correspondence should be addressed: Dept. of Pathology and Laboratory Medicine, University of Cincinnati, 231 Albert Sabin Way, Cincinnati, OH 45267-0529. Tel.: 513-558-3707; Fax: 513-558-2289; E-mail: Sean.Davidson@UC.edu.

[1] The abbreviations used are: HDL, high density lipoprotein; apoA-I, apolipoprotein A-I; DSP, dithiobis(succinimidyl propionate); DTT, dithiothreitol; HPLC, high pressure liquid chromatography; MS, mass spectroscopy; LCMS, liquid chromatography mass spectrometry; SPB, standard phosphate buffer; TOF, time of flight; TIC, total ion chromatogram; POPC, 1-palmitoyl 2-oleoyl phosphatidylcholine; rHDL, reconstituted HDL; DTT, dithiothreitol.

belt" model for a reconstituted HDL particle containing two molecules of apoA-I. In this model, two ring-shaped molecules of apoA-I are stacked on top of each other with both molecules forming an almost continuous helix that wraps around the perimeter of the phospholipid disc in an anti-parallel orientation. Computer analysis of the model predicted a particular registry between the monomers resulting in the greatest potential for salt bridge connections between the two molecules. An alternative belt-like model, initially suggested by Brouillette (15), predicts that apoA-I molecules are arranged in a hairpin orientation. In this model, about half of the molecule interacts with one leaflet, there is a turn, and the other half runs anti-parallel to the first on the opposing leaflet. This idea was supported by fluorescence energy transfer experiments performed by Tricerri *et al.* (16). The model preserves the potential for stabilizing salt bridge interactions between the same residues that were proposed for the double belt, although these must occur intramolecularly in the hairpin model. We have proposed a third model termed the "Z" belt orientation (13). This arrangement is similar to the hairpin except that, instead of traversing back along itself, the molecule proceeds in the same direction on the opposing leaflet, giving the potential for interlocking interactions between the molecules.

Although traditional spectroscopic methods such as fluorescence and circular dichroism have proven useful for studying the generalities of apoA-I structure in rHDL, data from these approaches are not suitable for high resolution modeling. Nuclear magnetic resonance (NMR) and x-ray crystallography data would be very useful, but these techniques have not yet been successfully applied to native rHDL particles. However, Bennett *et al.* (17) recently reported an elegant study that demonstrated the power of combining high precision mass spectrometry/peptide analysis with cross-linking chemistry to identify sites of interaction between two protein molecules. Young *et al.* (18) used a similar approach in combination with a sequence threading technique to generate a structure of monomeric human fibroblast growth factor that matched well with the known NMR structure of the protein. In this work, we report the successful adaptation of this approach to the problem the spatial relationships of two molecules of apoA-I on the edge of a discoidal HDL particle. The results provide the most comprehensive determinations of distance constraints within an rHDL particle and strongly confirm the presence of the salt bridge interactions predicted by Segrest *et al.* (14) present in both the double belt and hairpin models.

EXPERIMENTAL PROCEDURES

ApoA-I Purification—Purified human plasma apoA-I was obtained from human HDL ($1.21 < \text{density} > 1.062$ g/ml) isolated as reported (19). Briefly, HDL was freeze-dried and extracted with chloroform/methanol. The pellet was suspended in 10 mM Tris HCl with 6 M urea and applied to a Q-Sepharose column (XK 2.6/40, Amersham Biosciences) pre-equilibrated and eluted at 4 ml/min at room temperature in the same buffer. Fractions containing apoA-I as determined by SDS-PAGE electrophoresis were dialyzed into 5 mM ammonium bicarbonate buffer and freeze-dried. Proteins were solubilized in 3 M guanidine for 1 h and then dialyzed into standard phosphate buffer (SPB) (20 mM sodium phosphate, 0.15 M NaCl, pH. 7.8) prior to use in reconstitution experiments.

Preparation of rHDL Particles—Reconstituted HDL (rHDL) particles were prepared using 1-palmitoyl 2-oleoyl phosphatidylcholine (POPC) (Avanti Polar Lipids, Alabaster, AL) at lipid to protein molar ratios of 90:1 according to the method of Jonas (5). Lipids were dried under nitrogen and resuspended in SPB. Deoxycholate (Fisher, deoxycholate: lipid, 1.3:1, w/w) was added and incubated at 37 °C for 1.5 h with mild vortexing every 15 min. The protein was added and incubated a 37 °C for 1 h. The cholate was removed by dialysis against SPB (5 changes of 2 liters for at least 4 h each at 4 °C). The particles were analyzed on a non-denaturing, native polyacrylamide Phast gel (Amersham Biosciences, Piscataway, NJ) (20). Before cross-linking, the particles were passed down a Superdex 200 gel filtration column (Amersham Biosciences) to remove unreacted protein and lipid. Fractions corresponding to the 96-Å diameter complex were pooled and concentrated by filtration. The phosphorus method of Sokolof and Rothblat (21) and the Markwell modification of the Lowry assay (22) determined the final phospholipid and protein concentrations, respectively. The atomic phosphorus standard was obtained from Sigma (St, Louis, MO).

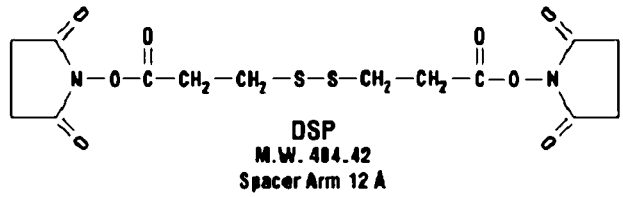

FIG. 1. **Chemical structure of dithiobis(succinimidyl proprionate) (DSP).**

Cross-linking, Reduction, and Generation of Tryptic Peptides—DSP (dithiobis(succinimidyl propionate)) (Pierce) was weighed out and dissolved in ice-cold Me_2SO to a concentration of 6.5 mg/ml and used within 5 min. A 7:1 molar ratio of DSP to apoA-I protein was added to a solution containing human apoA-I/POPC discs in SPB on ice at a concentration between 0.5 and 1.0 mg/ml. The reaction was incubated at 4 °C for 24 h with periodic vortexing. The reaction was quenched by adding a stock of 1 M Tris, pH.7.8, to a final Tris concentration of 100 mM. The samples were dialyzed into 5 mM ammonium bicarbonate to remove any unreacted cross-linker and were lyophilized to dryness. The lipids were extracted with chloroform/methanol, and the protein fraction was solubilized in SPB with 3 M guanidine HCl. In some experiments, the monomeric protein was separated from the dimeric protein after cross-linking by passage down a Superdex 200 column equilibrated in the same buffer. Fractions corresponding to the dimer and monomer were collected, concentrated, and dialyzed into SPB. Each cross-linked sample was split in two equal fractions. One was labeled "reduced" and incubated with 25 mM dithiothreitol (DTT) from a stock solution in water for 2 h at 37 °C. The other fraction was labeled "x-linked" and treated identically except receiving the same volume of water instead of DTT. Both samples were dialyzed against 5 mM ammonium bicarbonate (in separate containers). 5% (weight of trypsin to apoA-I) sequencing grade trypsin (Promega) was allowed to digest the protein at 37 °C for 2 h. The samples were lyophilized to dryness in a microcentrifuge tube in 100-μg aliquots. The samples were stored at −20 °C until used.

Mass Spectrometry—Liquid chromatography mass spectrometry (LCMS) experiments were carried out on a Sciex QSTAR DE (quadrupole time-of-flight (TOF)) mass spectrometer fitted with an atmospheric electrospray ionizer controlled by using Analyst QS 1.1 software (Applied Biosystems). The spectrometer was programmed for TOF-MS scans from 100 to 2800 atomic mass units at a 1.0-s accumulation time. An Agilent 1100 capillary HPLC with an Agilent ZORBAX SB-C18 0.5-mm × 15-cm reverse phase column at a flow rate of 7.5 μl per min was used to separate tryptic peptides prior to introduction into the mass spectrometer by complex peptide gradient chromatography. Lyophilized samples were solubilized in mobile phase A (distilled water with 0.1% trifluoroacetic acid) and eluted with a 0–100% gradient of mobile phase B (95% acetonitrile in water with 0.085% trifluoroacetic acid). 160 pmol of protein was injected per run. The mass spectra were internally calibrated by a three-point linear method based on the monoisotopic masses of the following peptides derived from apoA-I: 154–160 (mass of 780.4242 Da), 161–171 (1300.6412 Da), and 62–77 (1931.9265 Da). These peptides were used because they exhibited prominent peaks in all samples and covered much of the expected mass range.

Data Analysis—Using the Analyst QS software, individual mass spectra were generated for each peak in the total ion chromatograph (TIC) for each sample. Masses that exhibited an intensity of at least 25 detector counts were recorded and identified in terms of ion type (M+H, M+2H, etc.). Monoisotopic masses were determined by averaging the ion series for each mass. The resulting list of masses was analyzed by the software GPMAW (ChemSW, Inc.) to assign a putative amino acid sequence identity as either an unmodified peptide of apoA-I or one or more peptides containing one or more DSP modifications (see Table I). To identify cross-links occurring between two peptides, a spreadsheet was used to sum the masses of all possible combinations of peptides and one or more intervening cross-links. The resulting data base of theoretical cross-linked peptide masses was then searched for a given experi-

TABLE I
Possible peptide modifications resulting from DSP cross-linking

Type of modification[a]	Peptide component(s)	Mass addition due to DSP
		Da per event
Intrapeptide cross-link	1 peptide only	173.9809
Intramolecular cross-link	2+ peptides in same apoA-1 molecule separated by at least one trypsin cleavage site	173.9809
Intermolecular cross-link	2+ peptides on different molecules of apoA-1 on the same rHDL particle	173.9809
Hydrolyzed cross-link	1 peptide	191.9915
Reduced cross-link	1 peptide	87.9983

[a] See the text for a description of how cross-links were assigned as intra- or intermolecular. All masses are expressed as the monoisotopic mass.

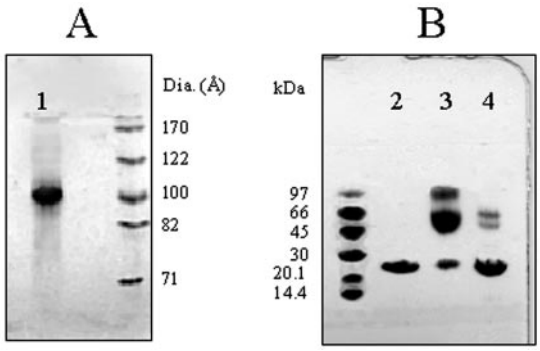

FIG. 2. **Characterization and cross-linking of a 96-Å apoA-I discoidal rHDL particle.** *A*, an 8–25% gradient native polyacrylamide Phast gel. *Lane 1*, a discoidal rHDL particle compared with a set of standards with the indicated hydrodynamic diameters. *B*, a denaturing 8–25% gradient SDS Phast gel of the same particle shown in *A*. *Lane 2*, an unmodified rHDL disc. *Lane 3*, DSP cross-linked rHDL. *Lane 4*, the same sample as in *lane 3* that has been reduced with DTT. All gels were stained with Coomassie Blue.

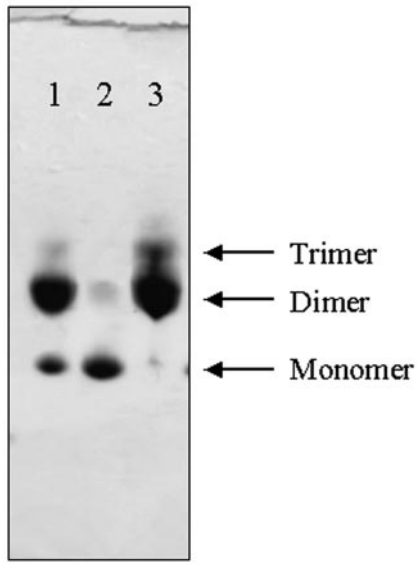

FIG. 3. **Gel filtration separation of the DSP cross-linked rHDL sample into monomeric and dimeric species.** The apoA-I/POPC rHDL shown in Fig. 2*A* was cross-linked with DSP as under "Experimental Procedures." After quenching the cross-linking reaction and removing unreacted cross-linker by dialysis, the cross-linked apoA-I was delipidated by chloroform:methanol extraction. The delipidated protein was applied to a Superdex 200 gel filtration column. Fractions containing the monomeric and dimeric forms were combined and concentrated by ultrafiltration. Shown is a denaturing 8–25% gradient SDS Phast gel stained with Coomassie Blue. *Lane 1* shows the cross-linked rHDL particle prior to separation. *Lane 2* shows the isolated monomeric form of the cross-linked protein. *Lane 3* shows the isolated dimeric form of the cross-linked protein.

mental mass. Mass identity assignments were made using the following criteria: 1) An assignment was made only if the experimental mass matched the theoretical mass within 40 ppm. This cutoff was sufficient to identify all single cleavage peptides in control experiments using unmodified apoA-I that had been completely trypsinized. 2) All peptides present in a cross-link must contain a Lys residue (not at the C terminus). 3) It was assumed that trypsin does not cleave on C-terminal side of a modified Lys residue (17). This feature of trypsin was advantageous because it reduced the complexity of the peptide mixture in the cross-linked sample. 4) Putative cross-links were assigned to masses present in the cross-linked sample but only if they completely disappeared when the sample was reduced with DTT. 5) Identities were assigned for putative intermolecular or intramolecular cross-links only if *all* peptide components were recovered with the appropriate number of reduced cross-links on eligible Lys residues after DTT reduction. 6) It was assumed that trypsin fully cleaved the protein at every opportunity, with no partials.

Distinguishing Intra- versus Intermolecular Cross-links—Intra- *versus* intermolecular cross-links were assigned by comparing the intensity of a given mass between spectra generated from dimeric and monomeric forms of apoA-I isolated from a cross-linked rHDL particle by gel filtration. The maximal detector count intensities for the entire ion series for a given mass were summed for the dimeric and monomeric cross-linked samples, respectively (23). The ratio of the dimeric/monomeric intensities was used to determine if a mass was substantially less prevalent in the monomeric *versus* the dimeric cross-linked protein. For a particular mass, an intensity ratio below 1.7 was identified as a putative intramolecular cross-link. By contrast, an intermolecular cross-link was proposed for ratios higher than 1.7. The appearance of small amounts of intermolecular cross-links in the monomeric sample was due to slight contamination of the dimeric form in the sample (see Fig. 3).

RESULTS

The Approach—The case of purified apoA-I on a well-defined rHDL particle with two molecules of apoA-I is essentially a homodimeric non-covalent interaction similar to that studied by Bennett *et al.* (17) using mass spectrometry. ApoA-I contains 21 Lys residues that are generally evenly spread throughout the molecule, making it a manageable candidate for this approach. The homodimer is first incubated with the homobifunctional cross-linking reagent DSP, which reacts with the ϵ-amine group of lysine residues. Numerous cross-links randomly form both intra- and intermolecularly depending on the number Lys residues within the reagent's spacer arm length of 12 Å. In addition, the cross-linker may bind to one Lys residue but fail to cross-link to a second Lys residue before spontaneous hydrolysis of the cross-linker (17). Once cross-linked, trypsin is used to cleave after Arg and Lys residues to generate a population of peptides, some of which are unmodified whereas others are cross-linked. An aliquot of the peptide mixture is treated with the reducing agent DTT to cleave the disulfide linkage within DSP (Fig. 1) to liberate any joined peptides. A second aliquot is left untreated. Both the cross-linked and reduced peptide mixtures are then separated by reverse phase HPLC and analyzed immediately upon elution by electrospray MS. Highly accurate mass spectra are taken for each peak as they elute from the column. If a particular mass is present in both chromatograms, it represents a peptide that was not modified by DSP. However, masses appearing in the cross-linked chromatogram, but not in the reduced chromatogram, indicate the presence of cross-linked peptides that were cleaved by DTT. Masses appearing in the reduced chromatogram are due to the

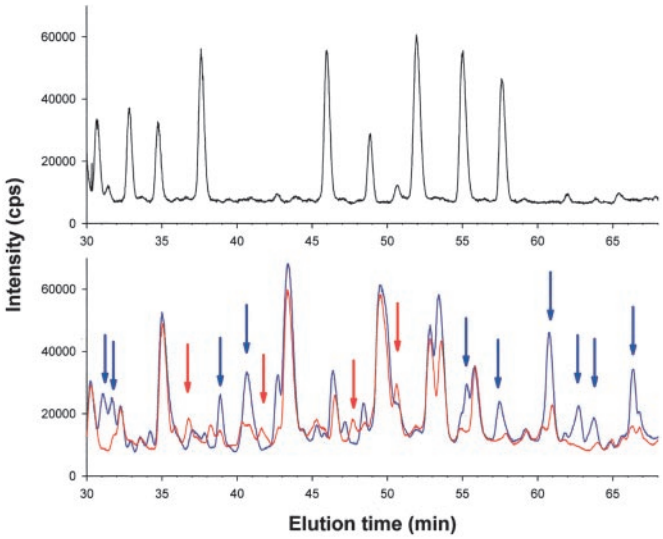

FIG. 4. **Total ion chromatographs (TIC) of untreated, cross-linked, and reduced samples of apoA-I in a 96-Å rHDL disc.** The peaks in each chromatogram correspond to peptides eluting from a reverse phase HPLC column. Each peak is a summation of the intensities of all ions striking the mass spectrometer detector at the indicated elution time. There are typically from one to five peptide masses present in each peak. *A*, TIC of an unmodified apoA-I rHDL sample. *B*, overlays of the TIC of the cross-linked dimer sample (*blue*) and the reduced monomer sample (*red*). The same mass of protein was injected for each run. Peaks that were unique to each chromatogram or exhibited a significant change in intensity are indicated with the *appropriately colored arrow*. The intensity is in detector counts per second.

newly freed peptides. The mass data are used to assign a sequence identity to each peptide by comparing the experimentally derived peptide mass to a data base of theoretical peptide masses generated from the known protein sequence and the known cleavage sites of trypsin. Unmodified peptides exhibit masses equal to the sum of the amino acids in the peptide sequence. The mass of a cross-linked peptide complex is the sum of each component peptide mass plus an intact cross-link (see Table I). Similarly, the mass of peptides containing cleaved cross-links is the sum of the peptide mass plus a reduced cross-link. After peptide identification, one can deduce that a cross-link was formed between peptides X and Y in the native dimer. If each peptide had a single Lys residue in the sequence, then one can conclude that the two Lys residues were within about 12 Å in the native protein structure.

rHDL Particle Reconstitution, Characterization, and Cross-linking Optimization—We generated a simple discoidal reconstituted rHDL particle with a diameter of 96 Å that is well known to contain two molecules of apoA-I and about 160 molecules of POPC (24, 25). This particle was selected because it is stable, easily produced in high yield *in vitro*, and has been used for computer simulation studies testing various models for apoA-I structure (14, 26, 27). Fig. 2*A* shows a non-denaturing PAGE analysis demonstrating the homogeneity of this particle. Its diameter measured 96 ± 3 Å, and it contained a final POPC to apoA-I ratio of 79 ± 4:1. To work out the conditions for cross-linking the apoA-I molecules on the complex with DSP, pilot experiments were performed in which the molar ratio of DSP to apoA-I was varied from 2.5:1 to 50:1 (data not shown). Ratios above 10:1 were sufficient to cross-link about 97% of the apoA-I to a dimeric form as visualized by SDS-PAGE. Higher ratios could drive the reaction completely to dimers. We chose a ratio of 7:1 for further experiments to minimize the chances of perturbing the particle structure and to give an opportunity to study the monomeric form of the cross-linked apoA-I (see below). Cross-linking at 4 °C was found to reduce the heterogeneity of the dimeric band *versus* incubations at room temperature probably by minimizing thermal motions within the particle. Fig. 2*B* shows the results of a typical cross-linking experiment under these conditions. ApoA-I (28 kDa) in the rHDL particle was cross-linked to a dimer (56 kDa), which could be mostly reduced back to a monomer by cleavage of the cross-link with DTT. We consistently observed a small amount of trimer (84 kDa) in these particles (Fig. 2, *lane 3*, *top band*) of about 5% of the total staining. We believe that this arose from a slight contamination of the 108-Å rHDL particles in our preparation, a complex with three molecules of apoA-I (24). A small percentage (about 5%) of the apoA-I remained as a monomer under these conditions. The appearance of the dimer upon DSP cross-linking was independent of the rHDL concentration between 0.5 to 5.0 mg/ml (data not shown), indicating that the cross-links were formed within rHDL particles and not between two rHDL particles (28). The presence of the cross-links did not change the average helical content of apoA-I from that of the unmodified form as measured by circular dichroism (both were about 75% helical), arguing that the cross-links did not significantly perturb the structure of the protein. Furthermore, we found that the ratio of cross-linker to apoA-I did not significantly affect the cross-links identifiable by MS (data not shown).

TABLE II
Peptide identification summary for the cross-linked dimer sample

Total number of masses found	110
Masses that were not affected by reduction	56
Masses containing DSP modification(s)	54
1 peptide[a] with hydrolyzed cross-link(s)	27
1 peptide with intrapeptide cross-link	4
2 or more peptides cross-linked	11
Predicted as intramolecular	4
Predicted as intermolecular	7
Currently unidentified[b]	12

[a] Identification of the experimentally derived masses were performed as described under "Experimental Procedures" using combinations of the known monoisotopic masses of tryptic peptides of human apoA-1 and the masses of the various forms of the cross-linker shown in Table I.

[b] Masses listed as "unidentified" were either masses for which we could not find a suitable combination of peptides or were not confirmable by the presence of all components in the reduced spectrum.

To get a sense of which experimental mass values originate from intermolecular *versus* intramolecular cross-links, a sample of cross-linked rHDL particles (*i.e.* the sample in Fig. 2, *lane 3*) was delipidated and then separated into the component monomeric and dimeric species by gel filtration chromatography (see "Discussion"). Fig. 3 shows that cross-linked apoA-I was resolved into two fractions, which will hereafter be referred to as the "cross-linked dimer" sample and the "cross-linked monomer" sample, respectively. Despite our best efforts, there was a slight contamination of <5% of the dimeric form in the cross-linked monomer fraction.

Mass Spectrometry—Four different delipidated apoA-I samples were prepared from rHDL particle preparations. They were the unmodified monomer, DSP cross-linked dimer, DSP cross-linked monomer, or cross-linked then reduced by DTT (reduced monomer). Each was subjected to trypsin digestion and the resulting peptides were analyzed by LCMS. Fig. 4*A* shows the total ion chromatograph (TIC) for the unmodified monomer. The TIC can be thought of in much the same way as a UV trace for a typical HPLC chromatogram except that, instead of an absorbance reading, the intensity value is generated by the mass spectrometer as a summation of all ions striking the detector during a scan at a particular elution time. The chromatogram shows some 30 peaks that contained a total of 37 masses. All but five of these masses were identifiable from

TABLE III
Identification of peptides with reduced cross-links after treatment of the cross-linked dimer with DTT to cleave DSP cross-links

Peptide[a]	Sequence[b]	Number of modifications	Theoretical mass	Observed mass
			Da	
239–243	**K**LNTQ	1	690.34	690.35
117–123	Q**K**VEPLR	1	956.51	956.52
207–215	A**K**PALEDLR	1	1099.57	1099.58
132–140	Q**K**LHELQEK	1	1239.63	1239.61
178–188	LEAL**K**ENGGAR	1	1244.62	1244.63
84–94	QEMS**K**DLEEVK	1	1422.64	1422.63
97–107	VQPYLDDFQ**K**K	1	1467.71	1467.67
107–116	**K**WQEEMELYR	1	1498.66	1498.67
11–23	V**K**DLATVYVDVLK	1	1549.84	1549.86
84–96	QEMS**K**DLEEV**K**AK	2	1709.77	1709.81
13–27	DLATVYVDVL**K**DSGR	1	1737.86	1737.88
46–61	LLDNWDSVTSTFS**K**LR	1	1968.96	1968.99
134–149	LHELQE**K**LSPLGEEMR	1	1995.98	1995.97
11–27	V**K**DLATVYVDVL**K**DSGR	2	2053.03	2053.04
28–45	DYVSQFEGSALG**K**QLNLK	1	2084.03	2084.09
189–206	LAEYHA**K**ATEHLSTLSEK	1	2115.03	2114.93
41–59	QLN**K**LLDNWDSVTSTFFSK	1	2296.14	2296.12
132–149	Q**K**LHELQE**K**LSPLGEEMR	2	2340.13	2340.17
196–215	ATEHLSTLSE**K**A**K**PALEDLR	2	2384.18	2384.23
216–238	QGLLPVLESF**K**VSFLASLEEYTK	1	2685.34	2685.37
62–83	EQLGPVTQEFWDNLE**K**ETEGLR	1	2705.27	2705.27
189–215	LAEYHA**K**ATEHLSTLSE**K**A**K**PALEDLR	3	3284.58	3284.42

[a] Identification of the experimentally derived masses was performed as described under "Experimental Procedures" using combinations of the known monoisotopic masses of tryptic peptides of human apoA-I and the masses of reduced cross-linker shown in Table I.
[b] Lysines containing reduced cross-links are in *boldface* type.

TABLE IV
Identification of cross-linked peptides in a 98Å rHDL particle cross-linked with DSP

Peptides involved		Lysines involved	Theoretical mass	Observed mass	TIC peak[a]	Intensity ratio[b]	Comment
			Da		*min*		
84–96		Lys^{88}-Lys^{94}	1707.75	1707.76	32.9	1.5	Intrapeptide
11–27		Lys^{12}-Lys^{23}	2051.01	2051.01	60.8	1.3	Intrapeptide
207–215x	207–215	Lys^{208}-Lys^{208}	2197.12	2197.13	40.6	3.3	Intermolecular
132–149		Lys^{133}-Lys^{140}	2338.11	2338.11	53.3	1.0	Intrapeptide
132–140x	132–140	Lys^{133}-Lys^{133}	2477.24	2477.25	29.5	2.9	Intermolecular
239–243x	28–45	Lys^{239}-Lys^{40}	2772.35	2772.37	72.9	0.3	Intramolecular
97–107x	97–107	Lys^{106}-Lys^{106}	2933.40	2933.44	45.3	2.7	Intermolecular
117–123x	134–149	Lys^{118}-Lys^{140}	2950.47	2950.48	42.7	3.9	Intermolecular
132–140x	134–149	Lys^{140}-Lys^{140}	3233.59	3233.58	40.6	3.8	Intermolecular
239–243x	216–238	Lys^{239}-Lys^{226}	3373.72	3373.75	69.9	2.8	Intermolecular
189–215	(+ H)	Lys^{195}-(Lys^{206} or Lys^{208})	3386.56	3386.52	51.9	1.4	Intrapeptide
207–215x	41–59	Lys^{208}-Lys^{45}	3393.69	3393.71	57.4	0.8	Intramolecular
117–123x	132–149 + H	Lys^{118}-(Lys^{140} or Lys^{133})	3398.62	3398.66	49.5	1.1	Intramolecular
117–123x	97–116 + H	Lys^{118}-(Lys^{106} or Lys^{107})	3878.75	3878.81	65.6	0.8	Intramolecular
189–206x	62–83	Lys^{195}-Lys^{77}	4818.28	4818.42	57.4	3.2	Intermolecular

[a] The TIC peak refers to the chromatographic peak in the total ion chromatogram (see Fig. 4*B*) in which a particular mass elutes from the HPLC column.
[b] The intensity ratio refers to the ratio of intensity of a given mass found in the cross-linked dimer *versus* the same mass found in the cross-linked monomer experiment. A ratio between 0.8 and 1.7 indicated little change in the intensity of the mass between the two samples predicting an intramolecular cross-link. Ratios higher than 1.7 were assigned as an intermolecular cross-link, because the cross-link was prevalent in the dimer sample. See "Experimental Procedures."

the theoretical cleavage of human apoA-I. The total apoA-I sequence coverage was 94% with all predicted peptides greater than two amino acids identified. There was no evidence of partial cleavage products generated by trypsin. The experimental monoisotopic mass values were all within 40 ppm of the theoretical value. Fig. 4*B* shows an overlay of the TIC of the cross-linked dimer (*blue*) and the DTT reduced monomer (*red*). It is clear that the peak pattern became more complex in cross-linked *versus* unmodified protein. Numerous differences between the cross-linked dimer and the reduced monomer chromatograms are apparent, including several peaks that completely disappear from the cross-linked dimer with new peaks appearing upon DTT treatment.

A list of masses was accumulated from individual mass spectra for each peak in all four chromatograms, and each was assigned an amino acid sequence identity using the criteria described under "Experimental Procedures." A summary of the identification process for masses observed in the cross-linked dimer spectrum is shown in Table II. Of the 110 masses found, 56 were insensitive to DTT reduction. Most of these were identified as peptides that either contained no Lys residues or contained Lys residues that had escaped modification by DSP. We found evidence for at least some degree of DSP modification for 19 of the 21 Lys in apoA-I. Lys^{28} and Lys^{59} appeared to be completely inaccessible to the cross-linker, suggesting that they may be buried within protein elements or lipid in the native rHDL particle. By contrast, lysines 40, 118, 133, 140, 226, and 239 were among the most active sites of modification, because they were commonly found with hydrolyzed cross-links or cross-links to numerous different residues. Although DSP can modify the free amino group on Asp-1 of lipid-free apoA-I, we saw no evidence that this residue was modified in the rHDL. We noted that the peaks with the highest intensity tended to be identified as those with hydrolyzed cross-links. This indicates that successful Lys-Lys cross-links were a relatively rare event when compared with the case when a cross-link forms at one

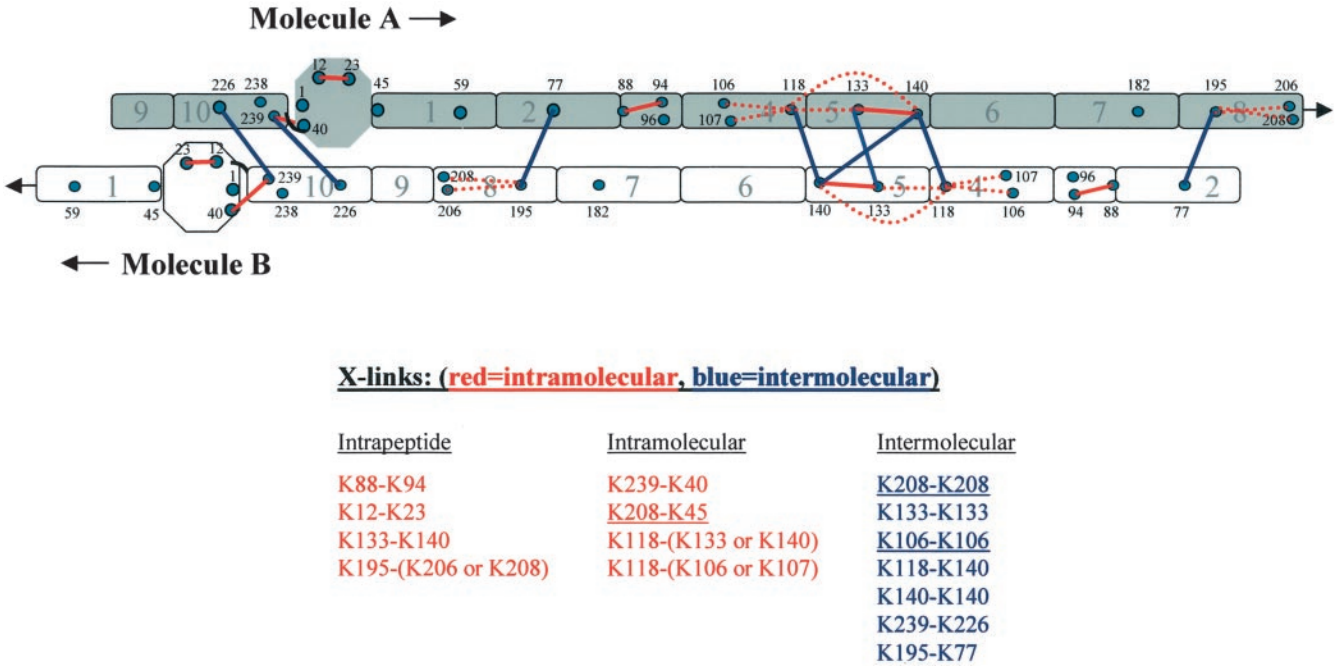

FIG. 5. **Graphical representation of the compatibility of the observed cross-links with the double belt model.** Two molecules of apoA-I are represented as either *gray* (molecule A) or *white* (molecule B) as if they had been taken off of the edge of a three-dimensional rHDL disc and laid flat on this printed page (or computer screen). The registry of the helices is shown in the 5/5 anti-parallel orientation as modeled by Segrest *et al.* (14). Residues 1–44 are predicted to exist in a globular conformation and are represented as an *octagon*. Each amphipathic helical segment is represented as a *rectangle* and is numbered according to the scheme of Roberts *et al.* (29). The 21 Lys residues are shown as *numbered green dots* in their approximate location along each molecule. Intramolecular and intrapeptide cross-links from Table IV are shown in *red*. Putative intermolecular cross-links are shown in *blue*. The Lys residues involved in each cross-link are listed on the figure using the same color scheme. Cross-links that *are not* consistent with the model are *underlined*. *Dotted lines* for some cross-links reflect cases when two Lys residues are in such close proximity that the cross-link may involve either residue.

site but cannot bind a second site before hydrolysis. Table III lists the peptides that were identified with reduced cross-links from the DTT-reduced monomer sample. The agreement between the observed monoisotopic mass and the theoretical masses illustrates the accuracy of the LCMS performed with this instrumentation. Table IV lists the identification of cross-links found in the cross-linked dimer sample and the dimer/monomer ratio of intensity of each mass.

DISCUSSION

In the present study, we successfully used LCMS and a thiol-cleavable cross-linker to derive the most comprehensive list of experimentally derived distance constraints available to date for apoA-I in a discoidal rHDL particle. Below, we discuss our interpretation of the MS data and then apply the information to three recently proposed models for apoA-I on a disc edge.

Data Interpretation—The random nature of the DSP cross-linking reaction provides the potential for a wealth of structural information, because many different Lys pairs residing within reach of the cross-linker spacer can be identified. However, in proteins with even modest numbers of Lys residues, the sheer number of potential combinations can quickly dilute the abundance of any one particular linkage. Therefore, the detection technique must be capable of accurately identifying small quantities of cross-linked peptides within a complicated background of hundreds of peptides that have large variations in abundance. For this reason, we chose electrospray LCMS *versus* the more commonly used matrix-assisted laser desorption ionization MS. The chromatography dimension of LCMS has the advantage of separating large numbers of peptides into manageable units prior to the analysis (23).

With over 110 experimentally derived masses in the cross-linked dimer sample, we set up a stringent set of criteria to faithfully identify cross-links (see "Experimental Procedures") and eliminate possible misidentifications. The accuracy of the instrumentation allowed the rejection of potential identifications if the theoretical mass differed from the experimental mass by greater than 40 ppm. This translates to a maximal error of about ± 0.20 Da for the largest peptide complex identified in this study (4818 Da) or about ± 0.08 Da for a 2000-Da peptide. With this degree of accuracy, most experimental masses had only one or perhaps two suitable identification possibilities. In addition, putative intermolecular cross-links were assigned only if *both* peptide components were observed with a reduced cross-link after DTT reduction. Our philosophy was that it was better to reject a "true" cross-link due to lack of evidence than to include a misidentified cross-link in the model analysis.

Another important issue when studying homodimer interactions is determining if experimentally observed cross-links occur inter- or intramolecularly. One cannot distinguish between the two by studying the covalently linked dimer alone (17) except in the case of two identical Lys residues cross-linked together. Therefore, we separated delipidated apoA-I from a cross-linked rHDL into the component monomeric and dimeric species by gel filtration. By definition, the dimer contains at least one pair of Lys resides cross-linked on two different molecules of apoA-I. However, because Lys residues are cross-linked randomly, it follows that a small fraction of molecules, by chance, would contain the spectrum of intramolecular cross-links but no intermolecular cross-links and thereby remain a monomer. The dimeric form also contains the spectrum of intramolecular cross-links but has the spectrum of intermolecular cross-links as well. By comparing the peptide maps from both samples, one can distinguish between the two types of cross-links. Because we had a small contaminant of dimer from our isolated cross-linked monomer sample, we used the ratio of

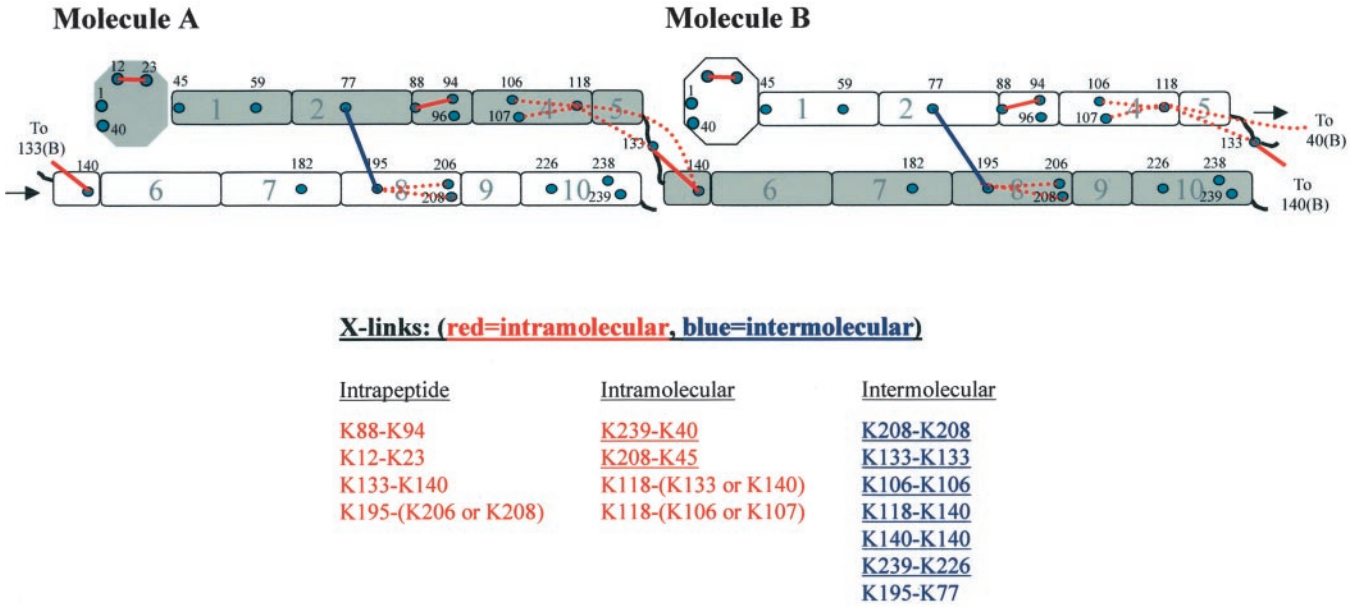

FIG. 6. **Graphical representation of the compatibility of the various cross-links with the parallel form of the "Z-belt" model.** The layout of the figure is as for Fig. 5.

the intensities of a given mass between the two samples as an index of their abundance in each. By analyzing unmodified peptides, we found that the ratio of the dimer/monomer intensity was between about 0.8 and 1.7; intramolecular cross-links are expected to fall in this range. Intermolecular cross-links, being more prevalent in the dimer sample, should give a ratio above 1.7. Note that all four peptides containing intrapeptide cross-links, *i.e.* two lysines cross-linked within the same tryptic peptide (and therefore must be intramolecular), were correctly indicated to be intramolecular cross-links as judged by the intensity ratio in Table IV. Conversely, all three cases of cross-links occurring between the same Lys residue (therefore must be intermolecular) were correctly identified.

Superimposing the Cross-links on Belt Models of ApoA-I—Figs. 5, 6, and 7 depict two-dimensional illustrations of three recently proposed models for apoA-I consistent with a predominantly helical belt orientation. We wish to make clear that no attempt was made to draw these models to any kind of molecular scale. Our purpose was to qualitatively test the general fit of the cross-links to the models. Detailed structural conclusions will await more sophisticated computer modeling studies currently under way. Fig. 5 shows the double belt model in its LL 5/5 orientation as proposed by Segrest *et al.* showing each Lys residue in its approximate position in apoA-I. For each model, we took the list of cross-links from Table IV and superimposed them onto the models. If the Lys residues were located within about one helix of each other as dictated by the model, they were drawn as an appropriately colored line. If a cross-link failed to easily fit the model, it was not drawn and the notation for the cross-link was underlined in the table under the figure. Fig. 5 shows that the majority of the cross-links fit the double belt model well. Because they are so close in the primary sequence, the four intrapeptide cross-links all fit the model. In fact, they fit all three models and therefore did not offer significant information on the spatial orientation of the two molecules. Three out of the four intramolecular cross-links also fit the double belt model. Of interest is the intramolecular cross-link between Lys^{239} and Lys^{40}. This could join the extreme C terminus with the globular region at the N terminus within the same molecule of apoA-I. The intramolecular cross-links involving Lys^{118} both fit, especially if Lys^{133} is the other Lys residue participating rather than Lys^{140}. Similarly, all but two of the putative intermolecular cross-links fit this model. Overall, of the 15 cross-links we identified, 12 were plausible in the double belt model.

Fig. 6 shows a model for apoA-I termed the "Z-belt" (named because of its shape on the edge of the disc), which we proposed to explain the presence of rHDL particles with three molecules on the disc edge (13). The model is attractive from a symmetry standpoint, because all molecules, including a third, can exist in the same conformation with helices in a belt orientation. Again, the intrapeptide cross-links all fit the model. However only two of the intramolecular and only one of the seven intermolecular cross-links were consistent. Fig. 7 shows two permutations of the helical hairpin in which the molecules are "head to head" and "head to tail." Comparison of this figure with Fig. 5 reveals that the hairpin models are related to the double belt in that all helical interactions are maintained with the same docking interface. The difference is that the helical interactions are intramolecular for the hairpins instead of intermolecular for the double belt. Despite this, all four intramolecular cross-links and four of the seven intermolecular cross-links fit the head-to-head model. However, none of the intermolecular links fit the head-to-tail version.

Comparing the total cross-link fits between the double belt and the head to head hairpin, both models allow 12 of the 15 cross-links. Surprisingly, they only differed by one cross-link each. The Lys^{208}-Lys^{45} intramolecular cross-link works in the hairpin but not in the double belt. By contrast the intermolecular Lys^{195}-Lys^{77} cross-link works in the double belt but not in the hairpin. All the other cross-links fit equally well to both models. The only cross-links that did not fit any of the models are the intermolecular Lys^{208}-Lys^{208} and the Lys^{106}-Lys^{106} connections. The reason for this is unclear, but an interesting speculation is that these cross-links represent an alternative registry of the double belt model. In Fig. 5, if one slides molecule A two helical positions to the right while keeping molecule B stationary (to a 4/4 orientation), both the Lys^{208}-Lys^{208} and Lys^{206}-Lys^{206} cross-links will fit the model, although all of the other cross-links that work in the 5/5 orientation will be broken. This may support the idea of variable helical registry that has been proposed by Li *et al.* (30), allowing for the possibility that a fraction of the particles are in a different registry. By contrast, we were unable to derive a permutation of the hairpin

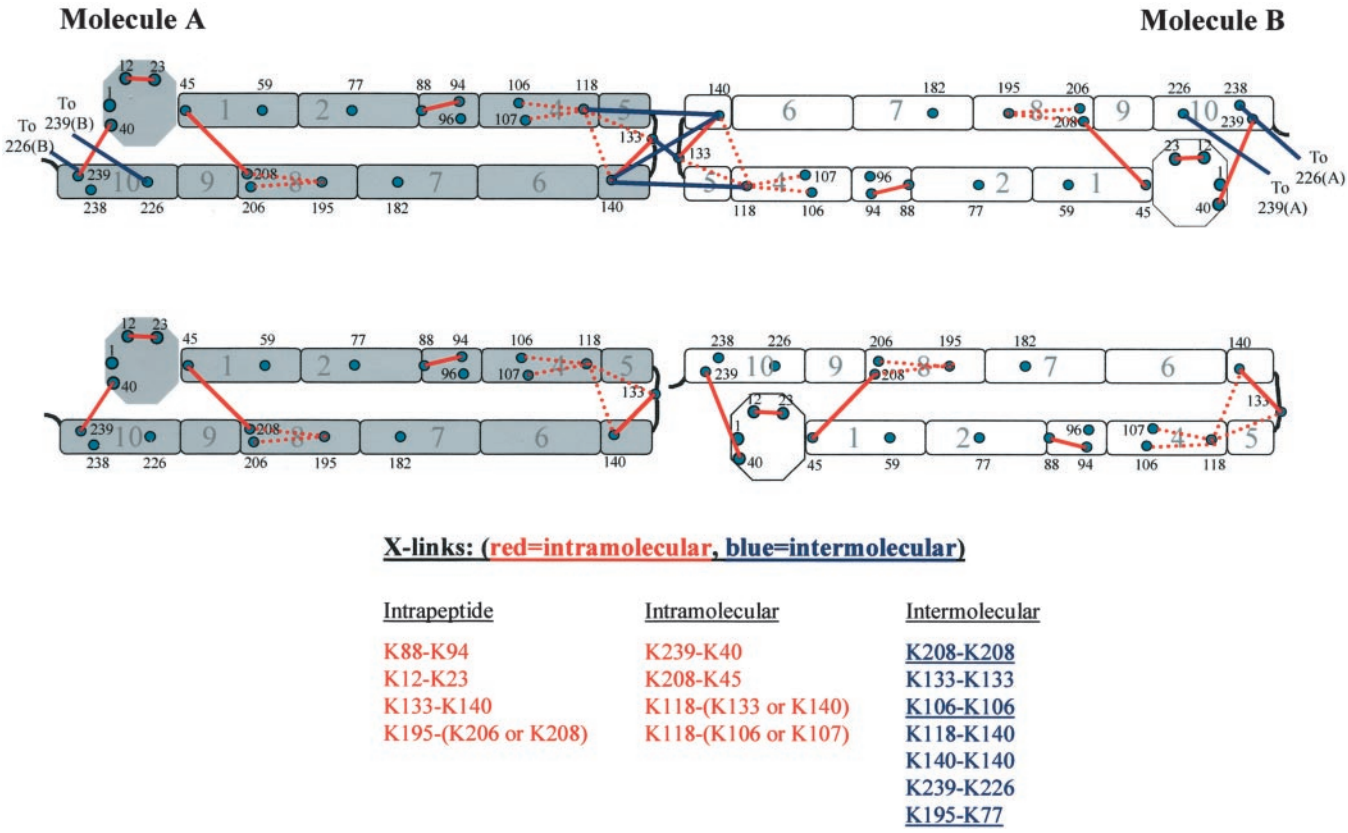

FIG. 7. **Graphical representation of the compatibility of the observed cross-links with the "head to head" (*top*) and "head to tail" (*bottom*) hairpin model combinations.** The layout of the figure is as for Fig. 5. Hairpins are shown with a *break* near the middle of the molecule in helix 5 (13).

models that could account for both of these cross-links.

From the model comparisons, it is clear that the cross-linking data strongly suggest that apoA-I exists in at least one state that maintains the salt-bridge docking interface proposed for the double belt model by Segrest *et al.* (14). This same interaction was also observed in an x-ray crystal structure of a lipid-free oligomer of apoA-I (10). Both the double belt and the hairpin models maintain this orientation, whereas the Z-belt does not. Thus, we feel that the Z-belt model can be safely ruled out. In terms of distinguishing between the double belt and the hairpin, we observed only two cross-links capable of distinguishing between the two models, the Lys^{208}-Lys^{45} (intramolecular) and the Lys^{195}-Lys^{77} (intermolecular). The fact that both of these cross-links were present may argue for a mixture of the two models existing in solution. Indeed, Segrest *et al.* (14) have suggested that rHDL particles containing three molecules of apoA-I may contain two present in the double belt conformation with the third adopting a hairpin motif to be accommodated on the disc edge. Because we did see a small amount of trimer formation within our rHDL particle preparation (see Fig. 2), one possibility is that we observed cross-links originating from particles containing two molecules of apoA-I both in the double belt conformation as well as some particles containing three molecules of apoA-I with two in the double belt and one in a hairpin. An alternative explanation is that both models may exist on different particles in solution. The particular orientation adopted by apoA-I may be determined by conformational factors at the time of initial lipid binding. If the hairpin does exist, either in conjunction with the double belt or alone on a particle, our data suggest that the head-to-head version is the best possibility as no evidence was obtained for the head-to-tail version.

We believe that this work represents a major step forward in bridging the gap between theoretical modeling and experimental evidence for apoA-I structure in its lipid bound state. Future work will focus on identifying additional distance constraints by using cross-linking agents with different spacer arm lengths, proteases with different specificities, and by improving our data analysis techniques to identify more complex cross-linked peptides. Furthermore, the approach has nearly boundless potential for applications in lipoprotein structural biology. Other forms of rHDL such as smaller particles, large ones containing three and four molecules of apoA-I, and even spherical particles can be compared by this method to identify flexible regions within apoA-I. In addition, because the masses that are derived from cross-links within apoA-I are detectible even in the presence of other apolipoproteins (18), the technique opens up the exciting possibility of studying the structure of apoA-I and other apolipoproteins in complex particles such as HDL isolated from human plasma.

REFERENCES

1. Xu, S., Laccotripe, M., Huang, X., Rigotti, A., Zannis, V. I., and Krieger, M. (1997) *J. Lipid Res.* **38,** 1289–1298
2. Rust, S., Rosier, M., Funke, H., Real, J., Amoura, Z., Piette, J. C., Deleuze, J. F., Brewer, H. B., Duverger, N., Denefle, P., and Assmann, G. (1999) *Nat. Genet.* **22,** 352–355
3. Bodzioch, M., Orso, E., Klucken, J., Langmann, T., Bottcher, A., Diederich, W., Drobnik, W., Barlage, S., Buchler, C., Porsch-Ozcurumez, M., Kaminski, W. E., Hahmann, H. W., Oette, K., Rothe, G., Aslanidis, C., Lackner, K. J., and Schmitz, G. (1999) *Nat. Genet.* **22,** 347–351
4. Brooks-Wilson, A., Marcil, M., Clee, S. M., Zhang, L. H., Roomp, K., van Dam, M., Yu, L., Brewer, C., Collins, J. A., Molhuizen, H. O., Loubser, O., Ouelette, B. F., Fichter, K., Ashbourne-Excoffon, K. J., Sensen, C. W., Scherer, S., Mott, S., Denis, M., Martindale, D., Frohlich, J., Morgan, K., Koop, B., Pimstone, S., Kastelein, J. J., and Hayden, M. R. (1999) *Nat. Genet.* **22,** 336–345
5. Jonas, A. (1986) *Methods Enzymol.* **128,** 553–582
6. Brouillette, C. G., Anantharamaiah, G. M., Engler, J. A., and Borhani, D. W. (2001) *Biochim. Biophys. Acta* **1531,** 4–46
7. Segrest, J. P. (1977) *Chem. Phys. Lipids* **18,** 7–22
8. Wlodawer, A., Segrest, J. P., Chung, B. H., Chiovetti, R. J., and Weinstein,

J. N. (1979) *FEBS Lett.* **104,** 231–235
9. Tall, A. R., Small, D. M., Deckelbaum, R. J., and Shipley, G. G. (1977) *J. Biol. Chem.* **252,** 4701–4711
10. Borhani, D. W., Rogers, D. P., Engler, J. A., and Brouillette, C. G. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 12291–12296
11. Koppaka, V., Silvestro, L., Engler, J. A., Brouillette, C. G., and Axelsen, P. H. (1999) *J. Biol. Chem.* **274,** 14541–14544
12. Panagotopulos, S. E., Horace, E. M., Maiorano, J. N., and Davidson, W. S. (2001) *J. Biol. Chem.* **276,** 42965–42970
13. Maiorano, J. N., and Davidson, W. S. (2000) *J. Biol. Chem.* **275,** 17374–17380
14. Segrest, J. P., Jones, M. K., Klon, A. E., Sheldahl, C. J., Hellinger, M., De Loof, H., and Harvey, S. C. (1999) *J. Biol. Chem.* **274,** 31755–31758
15. Brouillette, C. G., and Anantharamaiah, G. M. (1995) *Biochim. Biophys. Acta* **1256,** 103–129
16. Tricerri, M. A., Behling Agree, A. K., Sanchez, S. A., Bronski, J., and Jonas, A. (2001) *Biochemistry* **40,** 5065–5074
17. Bennett, K. L., Kussmann, M., Bjork, P., Godzwon, M., Mikkelsen, M., Sorensen, P., and Roepstorff, P. (2000) *Protein Sci.* **9,** 1503–1518
18. Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., and Dollinger, G. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 5802–5806
19. Lund-Katz, S., and Phillips, M. C. (1986) *Biochemistry* **25,** 1562–1568
20. Davidson, W. S., Gillotte, K. L., Lund-Katz, S., Johnson, W. J., Rothblat, G. H., and Phillips, M. C. (1995) *J. Biol. Chem.* **270,** 5882–5890
21. Sokoloff, L., and Rothblat, G. H. (1974) *Proc. Soc. Exp. Biol. Med.* **146,** 1166–1172
22. Markwell, M. A., Haas, S. M., Bieber, L. L., and Tolbert, N. E. (1978) *Anal. Biochem.* **87,** 206–210
23. Kapron, J. T., Hilliard, G. M., Lakins, J. N., Tenniswood, M. P., West, K. A., Carr, S. A., and Crabb, J. W. (1997) *Protein Sci.* **6,** 2120–2133
24. Jonas, A., Toohill, K. L., Krul, E. S., Wald, J. H., and Kezdy, K. E. (1990) *J. Biol. Chem.* **265,** 22123–22129
25. McGuire, K. A., Davidson, W. S., and Jonas, A. (1996) *J. Lipid Res.* **37,** 1519–1528
26. Phillips, J. C., Wriggers, W., Li, Z., Jonas, A., and Schulten, K. (1997) *Biophys. J.* **73,** 2337–2346
27. Brasseur, R., Lins, L., Vanloo, B., Ruysschaert, J. M., and Rosseneu, M. (1992) *Proteins* **13,** 246–257
28. Swaney, J. B., and O'Brien, K. (1978) *J. Biol. Chem.* **253,** 7069–7077
29. Roberts, L. M., Ray, M. J., Shih, T. W., Hayden, E., Reader, M. M., and Brouillette, C. G. (1997) *Biochemistry* **36,** 7615–7624
30. Li, H. H., Lyles, D. S., Pan, W., Alexander, E., Thomas, M. J., and Sorci-Thomas, M. G. (2002) *J. Biol. Chem.* **277,** 39093–39101