Proposed mechanisms for binding of apo[a] kringle type 9 to apo B-100 in human lipoprotein[a]

Juan Guevara, Jr.,* John Spurlino,‡ Amy Y. Jan,* Chao-yuh Yang,* Alexander Tulinsky,§ B. V. Venkataram Prasad,‡ John W. Gaubatz,* and Joel D. Morrisett*‡

Departments of *Medicine, and ‡Biochemistry, Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas 77030; and §Department of Chemistry, Michigan State University, East Lansing, Michigan 48824 USA

ABSTRACT The protein component of human lipoprotein[a] consists primarily of two apolipoproteins, apo[a] and apo B-100, linked through a cystine disulfide(s). In the amino acid sequence of apo[a], Cys4057 located within a plasminogen kringle 4-like repeat sequence (3991–4068) is believed to form a disulfide bond with a specific cysteine residue in apo B-100. Our fluorescence-labeling experiments and molecular modeling studies have provided evidence for possible interactions between this apo[a] kringle type and apo B-100. The fluorescent probe, fluorescein-5-maleimide, was used in parallel experiments to label free sulfhydryl moieties in lipoprotein[a] and low-density lipoprotein (LDL). In apo B-100 of LDL, Cys3734 was labeled with the probe, but this site was not labeled in autologous lipoprotein[a]. The result strongly implicates Cys3734 of apo B-100 as the residue forming the disulfide linkage with Cys4057 of apo[a]. To explore possible noncovalent interactions between apo B-100 and apo[a], the crystallographic coordinates for plasminogen kringle 4 were used to generate molecular models of the apo[a] kringle-repeat sequence (3991–4068, LPaK9), the only plasminogen kringle 4 type repeat in apo[a] having an extra cysteine residue not involved in an intramolecular disulfide bond. The Cys4057 residue (henceforth designated as Cys67 in the LPaK9 sequence) is believed to form an intermolecular disulfide bond with a cysteine of apo B-100. In computer graphics molecular models of LPaK9, Cys67 is located on the surface of the kringle near the lysine ligand binding site. Selected segments of the LDL apo B-100 sequence that contain free sulfhydryl cysteines were subjected to energy minimization and docking with the ligand binding site and adjacent regions of the LPaK9 model. In the docking experiments, apo B-100 segment 3732–3745 (PSCKLDFREIQIYK) displayed the best fit and the largest number of van der Waals contacts with models of LPaK9. Other apo B-100 peptides with sulfhydryl cysteine were found to be less compatible when minimized with this kringle. These results support and extend previously suggested mechanisms for a complex interaction between apo[a] and apo B-100 that involve more than a simple covalent disulfide bond.

INTRODUCTION

In humans, elevated plasma levels of lipoprotein[a] (Lp[a])[1] have been highly correlated with the incidence of coronary and cerebral artery disease (2–5). Lp[a] consists of two major and distinctly different high molecular weight apoproteins (6–8). Apoprotein[a] (apo[a]) is a highly glycosylated, hydrophilic protein (7–13) with little apparent affinity for lipid. It is disulfide-linked to one molecule of apolipoprotein B-100 (apo B-100) (7, 10, 14), a highly hydrophobic apolipoprotein that is partly imbedded in a lipid-rich, pseudomicellar particle. Apo[a] has been shown to be colocalized with apo B-100 in aorto-coronary bypass vein grafts (15) and the aorta (16), directly implicating Lp[a] in atherosclerosis. However, neither elucidation of a positive role for Lp[a] nor definition of the mechanism(s) by which Lp[a] leads to atherosclerosis have been reported.

The single apo[a] polymorph whose primary structure has been inferred by cDNA sequencing (1) has a mass of ~530 kD and consists of two distinct domains: a kringle domain and a serine protease domain. There are 38 repeated sequences in the kringle domain of this apo[a] polymorph. Careful examination of the amino acid sequence of each repeat suggests 11 different kringle types (17). Kringle types 1 through 10 are similar but not identical to plasminogen kringle 4 (PGK4). The kringle type 2 repeat sequence (LPaK2) is serially copied 28-fold. The residues in LPaK9 and PGK4 are numbered 1 through 78 unlike the system used in reference 27 for PGK4 that is based on the prothrombin kringle 1 sequence of 80 residues and, therefore, contains two deletions. Although apo[a] kringle types 1–10 and PGK4 contain 78 residues, there exist kringle sequences in plasminogen and other proteins that contain >80 residues. For a review, see Table 1 of reference 39. A kringle is a highly conserved, tri-loop structure stabilized by three disulfide bridges; its function is noncatalytic but rather involves recognition and binding (18, 19). Several hemostasis proteins contain the kringle motifs (20, 21). The structures of kringles in apo[a] may provide insight into the physiologic role(s) of the Lp[a] particle and help define the mechanism(s) by which the lipoprotein expresses its atherogenicity.

The three kringle structures elucidated thus far by x-ray crystallography show a conservation of the overall structure (22–29). These observations provide a basis for using the three-dimensional coordinates of these structures for evaluating other kringle-like sequences, es-

Address correspondence to Dr. Joel D. Morrisett, The Methodist Hospital, MS A601, 6565 Fannin Street, Houston, TX 77030.

[1] *Abbreviations used in this article:* apo[a], apoprotein[a]; apo B-100, apolipoprotein B-100; EDTA, ethylenediaminetetraacetate; F5M, fluorescein-5-maleimide; HPLC, high-performance liquid chromatography; LDL, low-density lipoprotein; Lp[a], lipoprotein[a]; LPaK9, lipoprotein[a] kringle type 9 (kringle repeat #36; see reference 1); PGK4, plasminogen kringle 4; TFA, trifluoroacetic acid; Tris, tris(hydroxymethyl)-aminomethane.

0006-3495/93/03/686/15 $2.00

pecially those of apo[a] that have a high degree of homology with PGK4. Therefore, we have generated models of LPaK9 based on the coordinates of PGK4 (27) and PGK4 liganded with ε-amino caproic acid (28). Unlike other apo[a] kringle types that contain six cysteine residues, LPaK9 contains a seventh cysteine that occurs at position 67 and is a good candidate for the site of disulfide linkage to apo B-100. However, the interactions between apo[a] and apo B-100 most likely also include noncovalent binding through one or more of the apo[a] kringles (30, 31). These interactions may involve ligands such as L-proline and, to a lesser extent, L-lysine (32). LPaK9 is an ideal region of apo[a] for understanding the interaction of apo[a] with apo B-100. LPaK9 is ~80% homologous to PGK4 and appears to have an intact lysine ligand-binding site, although some of the nonhomologous replacements are in the ligand binding region. For example, Asp57, an essential amino acid for the binding of lysine to PGK4, is replaced with glycine in LPaK9. Furthermore, Trp32 and Arg35 in LPaK9, which replace, respectively, Arg32 and Lys35 in PGK4, influence the replication of the PGK4 ligand-binding site by their overall mass and charge characteristics (Fig. 1). The effect of these amino acid replacements in critical positions of the kringle structure can be evaluated by molecular modeling.

In this study, molecular models of LPaK9 have been used to explain chemical data that identify regions of apo B-100 that may form a disulfide bond with apo[a]. Sixteen of 25 cysteine residues in apo B-100 occur in disulfide form (33), with possibly 7 other cysteine residues in sulfhydryl form on the surface of the particle. Our approach was to label the free thiol cysteine residues of low-density lipoprotein (LDL) and Lp[a] with a sulfhydryl-specific fluorescent alkylating reagent. Trypsin-digest fragments of apo B-100 from both LDL and Lp[a−], which contained fluorescent label, were then compared by high-performance liquid chromatography (HPLC). LDL apo B-100 labeled fragments that were not duplicated in similar sequence fragments of Lp[a] apo B-100 fragments then indicated the location of the intermolecular disulfide bond. Molecular modeling was used to determine the structural complementarity between fragments of apo B-100 and the folded LPaK9 structure. Amino acid ligands and trypsin-releasable apo B-100 polypeptide fragments that contained a free thiol cysteine residue (33, 34) were modeled along with LPaK9. In this report we present results of the fluorescent labeling experiments with Lp[a] and LDL, together with results from molecular modeling and docking experiments that suggest the existence of important noncovalent interactions between apo[a] and apo B-100.

MATERIALS AND METHODS

Materials

HPLC grade solvents were obtained from J. T. Baker (Philipsburg, New Jersey). Spherisorb ODS II and Hypersil ODS HPLC columns were purchased from Phenomenex (Rancho Palos Verdes, California). Phenylmethylsulfonyl fluoride, sodium azide, tris-(hydroxymethyl)-aminomethane (Tris), sodium sulfate, 6-aminohexanoic acid (ε-aminocaproic acid), glycerol, and ethylenediaminetetraacetate (EDTA) were from Sigma Chemical Company (St. Louis, Missouri). Aprotinin (Trasylol) was from Miles Inc., FBA Pharmaceuticals (West Haven, Connecticut). Lysine-Sepharose was from Pharmacia LKB (Alameda, California). Fluorescein-5-maleimide (F5M) were from Molecular Probes (Eugene, Oregon). Tosylphenylalanyl-chloromethyl ketone-treated trypsin was from Worthington Biochemicals (Freehold, New Jersey). Reagents and solvents for the automatic gas-phase sequencer were products of Applied Biosystems Inc. (Foster City, California). All other chemicals used were of the highest quality available.

Preparation of Lp[a] and LDL

Purified Lp[a] and LDL were prepared by methods described elsewhere (8). Lp[a] was isolated from the 1.063–1.130 density fraction by lysine-Sepharose chromatography. This fraction, in 50 mM phosphate buffer (pH 7.5), was applied to the affinity column that was then washed with the same buffer until the optical density was below 0.05. The bound Lp[a] was then eluted with 100 mM 6-aminohexanoic acid and 500 mM sodium chloride in 50 mM phosphate buffer (pH 7.5). LDL was prepared from the 1.020–1.063 density fraction using lysine-Sepharose affinity chromatography to remove any contaminating Lp[a].

Fluorescein labeling of lipoprotein particles

Free sulfhydryl groups in Lp[a] and LDL were labeled with a fluorescent probe using a slight modification of the method described previously (35). Briefly, to 20.0 mg of Lp[a] (in ~20 ml) or LDL (concentration was usually >2.0 mg/ml) was added 2.5 ml 1.0 M Tris-acetate and 10 mM EDTA, pH 7.6, 0.35 g sodium sulfate, 32.2 mg 6-aminohexanoic acid, 2.5 ml glycerol, and 7.2 mg F5M. The total volume of the reaction mixture was adjusted to 25 ml with nanopure grade water. The samples were gently agitated in darkness at 30°C for 16 h. Unreacted F5M was removed using a Sephadex G25 (fine pore) column (2.5 cm ID × 40 cm h) that had been equilibrated with several volumes of the Tris-acetate-EDTA buffer. The labeled lipoprotein peak eluted from the column in the void volume and was collected as a single fraction. Each preparation was made 20.0 mM dithiothreitol and 10.0 mM 6-aminohexanoic acid and reduction was allowed to proceed at ambient temperature for 3 h. Potassium bromide was then added to each sample to obtain a solution density of 1.120 for separating the labeled LDL and labeled Lp[a−] particles from nonlipoproteins. Centrifugation was performed using an ultracentrifuge (model L8-80; Beckman Instruments, Fullerton, California) and a 50.2 Ti rotor at 40,000 rpm for 22 h.

The labeled LDL and labeled Lp[a−] that floated to the top of the centrifuge tubes were collected and dialyzed against 6 liters of nanopure water with at least three changes. Pellets of higher density protein such as apo[a] were also collected, resuspended in nanopure water, and dialyzed as before. Each sample was then shell frozen and lyophilized. Lyophilized powders were delipidated using an ethanol:diethyl-ether (1:3) solvent mixture as described previously (8). No fluorescence was detectable in the apo[a] samples after delipidation.

Tryptic digestion

Delipidated apo B-100 from LDL and from Lp[a−] were hydrolyzed using L-1-*p*-tosylamino-2-phenylethyl chloromethyl ketone-treated trypsin as described by Coleman et al. (35).

Polypeptide isolation

Tryptic polypeptides were separated by HPLC as described earlier (34). Briefly, aliquots of the digest mixture were first applied to a VY-

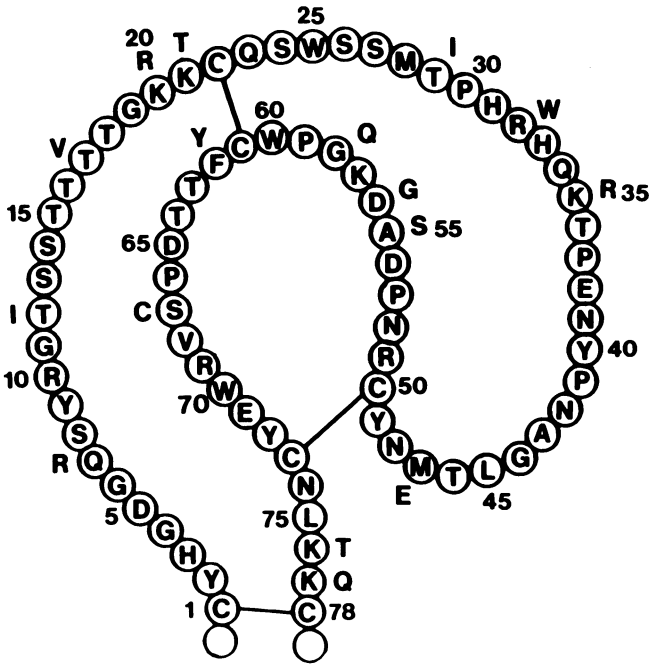

FIGURE 1 Sequence comparison of human plasminogen kringle 4 and apo[a] kringle type 9. The 16 residue replacements in the apo[a] kringle type 9 are shown outside the circled PGK4 residue.

DAC (Hesperia, California) C_4 6.25 mm ID × 25 cm 1 reverse-phase column. Peptides were eluted with a linear 1% per minute gradient of a trifluoroacetic acid buffer system (A, 0.1% trifluoroacetic acid [TFA] in water; B, 0.08% TFA in 95% acetonitrile and 5% water). Each fluorescent polypeptide peak was collected and then applied to a Phenomenex 5 μM Hypersil ODS 4.6 mm ID × 25 cm 1 column and eluted with a sodium phosphate buffer, pH 6.0 (solvent A), and acetonitrile:water, 9:1, (solvent B) system linear gradient (34). Further purification of the fluorescent peptide was accomplished by rechromatographing the isolated peak on the Phenomenex column using the same eluting system, except that solvent A was at pH 4.0. All chromatography was conducted at 50°C. Ultraviolet adsorption was monitored at 220 nm and fluorescence at 515 nm using Applied Biosystems Inc. detectors.

Sequence analysis of fluorescent peptides

Sequence analysis of the fluorescent peptides was performed as described by Coleman et al. (35) using an Applied Biosystems Inc. automated gas-phase sequencer.

Molecular modeling systems and methods

Modeling was performed on molecular graphics workstations from Silicon Graphics Inc. (Mountain View, CA) and Evans and Sutherland, Inc. (Salt Lake City, UT) using the crystallographic modeling programs CHAIN (36), Polygen Corporation's (Waltham, MA) QUANTA 3.0, and the Adopted Basis Newton-Raphson energy minimization procedure in CHARMm (version 21). There are 16 amino acid substitutions in LPaK9 that distinguish this apo[a] kringle from PGK4 (Fig. 1). Five of these substitutions are also present in nine other apo[a] kringle types that are similar to PGK4. The five substitutions that are conserved in PGK4-like kringles of apo[a] are Val17, Arg20, Thr21, Arg35, and Thr76. There are 11 additional residue replacements in LPaK9 that may lead to slight structural differences from PGK4. In PGK4, Lys20-Lys21 and Lys76-Lys77 of PGK4 are replaced with arginine-threonine and threonine-glutamine in LPaK9. In PGK4, Thr29 and Arg32 are changed to isoleucine and tryptophan in LPaK9; these two residues occur on the surface of the kringle. In addition, Lys35, Ala55, Asp56, and Gly58 of the ligand binding region of PGK4 are altered to arginine, serine, glycine, and glutamine in LPaK9.

The effect of the above changes on the kringle structure was studied using several different approaches to produce models of LPaK9. One approach involved adding each amino acid residue at the appropriate alpha-carbon coordinates to the x-ray crystallographic structure of PGK4; the amino acid side groups were added without positional bias, i.e., chi torsion angles were the default all-*trans* form of the Alchemy II program (Tripos Associates, St. Louis, MO). The alpha-carbon atoms were held fixed as the model was then subjected to energy minimization. All the atoms were then allowed to move, and the molecule was refined further. This model is henceforth referred to as the unbiased LPaK9 model. A second approach was to use the complete PGK4 crystallographic structure and make the 16 amino acid substitutions required to convert the molecule to LPaK9 (Fig. 1). The structure of the molecule was then refined as previously described. This model is referred to as the biased LPaK9 model. Typically, amino acids such as tryptophan, tyrosine, and phenylalanine were regularized to maintain strict aromatic ring planarity within CHAIN (36) before the final energy minimization was applied. Although other models of LPaK9 based on PGK4 were studied, only these two models of LPaK9 (biased and unbiased) are described in the present report.

Amino acids phenylalanine, proline, leucine, and arginine were minimized positioned in the ligand-binding site of the biased LPaK9 model. Trypsin-releasable LDL apo B-100 sequences (34) containing cysteine residues that were selected for ligand docking experiments with the biased LPaK9 model are listed in Table 1. Apo B-100 fragments were constructed using the programs Sequence Builder and Protein Design (Polygen Corp.). Several different approaches were used in modeling fragments of apo B-100 with the biased model of LPaK9. In one approach, fragments of eight amino acids, with the cysteine in position 3, were constructed for determining the sequence that formed the most contacts with residues adjacent to LPaK9 Cys67, residues between Cys67 and the ligand-binding site, and residues within the binding site (i.e., Arg35, Asp54, Gly56, Trp60, Arg69, Trp70, and Tyr72). In a second approach, the previous method was repeated but with a distance constraint of 2.0 Å imposed for a disulfide bond between the LPaK9 Cys67 and the apo B-100 fragment. This complex was then minimized with the kringle held fixed. After several cycles of minimization, the distance constraint was removed and the complex was subjected to further minimization. Finally, the kringle itself was released and one cycle of minimization was performed.

A third approach was intuitive and based on the results obtained through the fluorescent labeling experiments. Apo B-100 fragment 3732–3746 (Table 1) of 15 amino acids was generated and contoured on the surface of the kringle with Phe3738 placed in the ligand-binding

TABLE 1 Polypeptide fragments from apo B-100 containing unbound cysteine residues used in molecular docking experiments with models of LPaK9

$^{NH}K_{1073}$ITEVALMGHLSCDT	L_{1083}SCDTKEERKIKGVI
T_{1383}TYDHKNTFTLSCDG	L_{1393}SCDGSLRHKFLDSN
S_{1466}SFYAKGTYGLSCQR	L_{1476}SCQRDPNTGRLNGE
D_{1623}GISTSATTNLKCSL	L_{1633}KCSLLVLENELNAE
F_{3722}HVPFTDLQVPSCKL	P_{3732}SCKLDFREIQIYKK
K_{3878}ADYVETVLDSTCSS	S_{3888}TCSSTVQFLEYELN
P_{4178}GKPGIYTREELCTM	E_{4188}LCTMFIREVGTVLS

^{NH}, direction of amino terminus. Underlined residues were placed in the ligand-binding site of LPaK9 with the residue nearest the free cysteine in the Trp60/Trp70 region.

site. Contouring or shaping of the fragment was accomplished by torsioning bonds and minimizing the fragment before placing it on the surface of the kringle. The cysteine of the fragment was placed proximal to Cys67 of LPaK9 without constraints, and the resulting complex was subjected to minimization. An additional protocol was used for apo B-100 fragments containing Cys3734 and Cys3890 to determine their complementarity with another region of the biased LPaK9 kringle. In this case, dipeptides were added to the eight amino acid fragments generated in the above constructions starting at the ligand-binding site (i.e., $Ile_{3741}Gln$ to Glu_{3740} and $Val_{3842}Glu$ to Thr_{3844}). Several minimization cycles were then performed after each addition.

RESULTS

Indirect chemical evidence of the LPaK9 disulfide linkage to apo B-100

Previously, Coleman et al. (35) identified two free cysteine residues in LDL apo B-100 at positions 3734 and 4190 using the fluorescent probe, 5-iodoacetamidofluoresceine. In the present study, Lp[a] and LDL were treated with F5M, a thiol specific reagent, to determine the difference in number and location of the free sulfhydryl groups in apo B-100 for each lipoprotein. The probe, F5M, was selected for these experiments because it is water soluble, thiol specific, and easily detectable. Although there are slight differences between the chromatograms of tryptic peptides from LDL apo B-100 and Lp[a−] apo B-100 detected at 220 nm, the elution profiles shown in Fig. 2, *A* and *B* share many common peaks and both are highly reproducible. In LDL, an apo B-100 fragment containing Cys3734 was labeled with F5M as had been shown previously (35). Apo B-100 fragments containing Cys4190, shown as fluorescent peaks 30–33, were weakly labeled in both LDL and Lp[a−]. Fluorescent peaks 1–29 (Fig. 2, *C* and *D*) were previously shown not to contain a cysteine residue (35). Although the peak patterns as determined by absorbance at 220 nm are similar in the two peptide chromatograms (Fig. 2, *A* and *B*), there are some apparent differences in the fluorescence chromatograms (Fig. 2, *C* and *D*). The Lp[a−] fluorescence chromatogram (Fig. 2 *D*) shows different labeling of noncysteine containing peptides, i.e., different labeling of other nucleophilic amino acids such as lysine. Fluorescent peaks 35 and 36 shown in Fig. 2 *C* contain the apo B-100 tryptic peptide cleaved at Lys3735 and include the sequence $L_{3710}NDLNSVLVMPTFHVPFTDLQVPSCK_{3735}$. The apo B-100 Cys3734 is the labeled residue in this tryptic peptide. These results are identical to those reported previously for apo B-100 in LDL from a different donor (35). Similar A_{220nm} absorbance patterns were obtained for tryptic peptides from apo B-100 in Lp[a−] eluted from the column (Fig. 2 *B*). In contrast, fluorescent peaks 35 and 36 were prominent in the LDL apo B-100 fluorescence chromatogram (Fig. 2 *C*) but were barely discernible in the Lp[a−] apo B-100 fluorescence chromatogram (Fig. 2 *D*). These results indicate that Cys3734 of apo B-100 in Lp[a] is involved in a disulfide linkage.

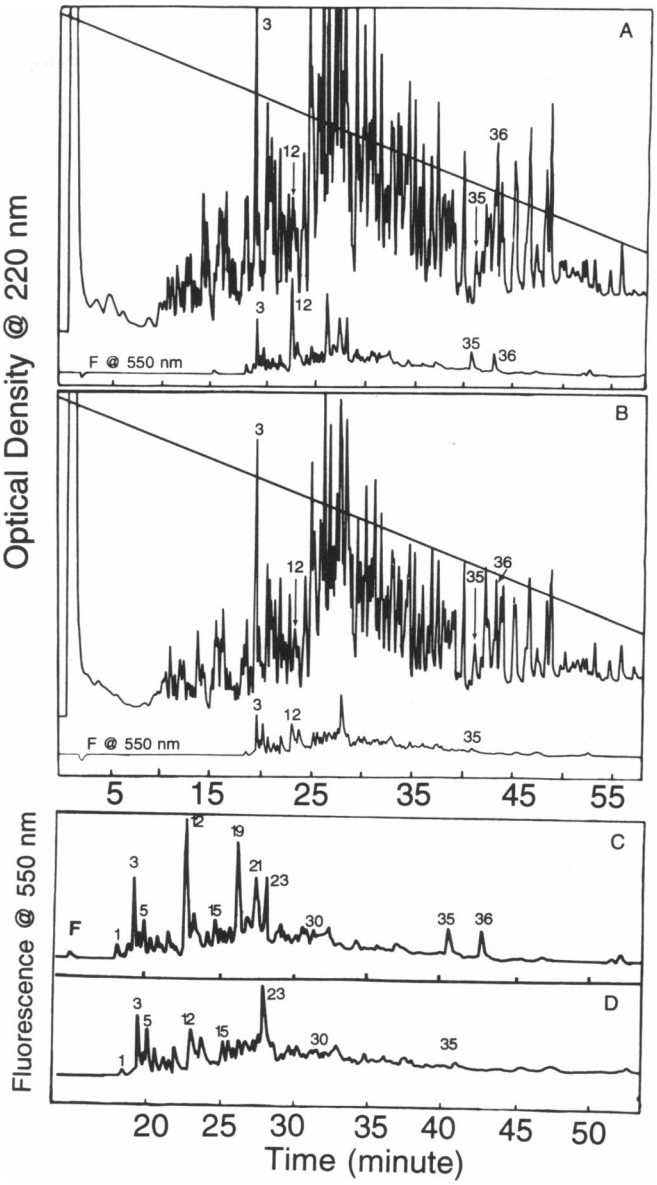

FIGURE 2 HPLC chromatograms of the tryptic peptides of F5M labeled apo B-100 from LDL (*A* and *C*) and from Lp[a−] (*B* and *D*). Primary separation of tryptic peptides was performed on a Vydac C_4 column (6.25 × 250 mm) using a TFA buffer system. Previous amino acid analysis of peaks 1 through 29 showed these to be fluorescent peptide or tryptic peptides that did not contain a Cys residue (35). Fluorescent peaks 35 and 36 are tryptic peptides resulting from apo B-100 cleavage at Lys3735 and both peptides contain Cys3734. Cys4190 is present in fragments that eluted as peaks 30 through 33; Cys4190 was labeled weakly in apo B-100 from both LDL and Lp[a].

Molecular modeling of LPaK9 and apo B-100 interactions

We considered it plausible that Cys3734 of apo B-100 forms a disulfide linkage with Cys67 of LPaK9 to form an Lp[a] particle. Therefore, molecular models of

TABLE 2 **CHARMm potential energy (kcal/mol) calculations for PGK4 and minimized molecular models of LPaK9**

	PGK4	LPaK9	
		Biased	Unbiased
Total E	−2,516	−2,538	−2,486
Bond E	23	22	25
Angle E	184	114	150
Dihedral E	260	215	279
Improper E	41	38	44
van der Waals E	−194	−222	−195
Electrostatic E	−2,831	−2,697	−2,788

LPaK9 were generated by making the appropriate amino acid side chain substitutions onto the PGK4 polypeptide backbone (unbiased model) as well as on the complete PGK4 structure (biased model). Before performing the docking experiments with apo B-100 fragments, each LPaK9 model was evaluated to determine the influence(s) the 16 amino acid alterations had on the structure of the complete kringle and its ligand binding site.

Modeling of LPaK9

Each LPaK9 model was subjected to several cycles of energy minimization until values comparable with those of PGK4 were obtained. Energy values shown in Table 2 for both LPaK9 models strongly support the probability that this segment of apo[a] exists in a kringle conformation. Comparisons of the C-alpha trace of each LPaK9 model with that of PGK4 were used to identify regions of the kringle where amino acid alterations cause a shift in the main chain. A schematic comparison of the C-alpha trace structure of PGK4 with those of the two models of LPaK9 is shown in Fig. 3. Although some shifts are seen between the biased LPaK9 model and PGK4, the unbiased LPaK9 shows changes throughout the main chain of the kringle. Shifts in the main chain appear to occur in both LPaK9 models at the putative ligand-binding region of the kringle. The largest shifts occur in the main chain of the unbiased LPaK9 model (purple) in the regions adjacent to positions 8, 17, 57, and 68. These main chain changes are at or near sites of amino acid alterations. The shifts are more pronounced in the main chain of the unbiased LPaK9 model and probably reflect the lesser degree of constraints imposed in building that model. However, residues that form the hydrophobic clusters and are known to stabilize the structure in PGK4 (27) are highly conserved in LPaK9 and must impart stability in a similar manner. The rms deviations of atoms of the main chain and aromatic and proline side groups were determined from a comparison of the LPaK9 models with PGK4 (Table 3). The values indicate that positional deviations of the side groups are again greater between the unbiased LPaK9 model and PGK4 than for those of the biased LPaK9 model and PGK4. These results show a change in the nature of

FIGURE 3 Schematic comparison of the C-alpha trace structure of PGK4 with LPaK9 molecular models. The C-alpha trace of the unbiased LPaK9 model (*purple*) is superimposed on the trace structure of PGK4 (*green*). The C-alpha trace of unbiased LPaK9 model was generated by making residue side group additions to the PGK4 mainchain. Prominent shifts of the polypeptide main chain are indicated for this LPaK9 model in the putative ligand-binding region at positions 8, 17, 57, and 68 and to a lesser degree at other sites in the kringle. The C-alpha trace of the biased LPaK9 model is shown in yellow. Small positional deviations are observed throughout the mainchain of this model as well.

overall hydrophobic core of the unbiased kringle structure that may, along with dihedral angles, indicate a less stable conformation. The Ramachandran angles of the unbiased and biased models of LPaK9 (Fig. 4, *A* and *B*) also suggest this possibility. The phi/psi angle distribution for the biased LPaK9 indicates that this model conforms well to allowed regions of the plot, whereas the unbiased LPaK9 has several residues in unfavorable regions. Several main chain changes in the biased LPaK9

TABLE 3 **Comparison of rms deviations of LPaK9 models from PGK4 hydrophobic core residues**

Residue position	Amino acid	Main chain rms (Å)		Side group rms (Å)	
		Unbiased	Biased	Unbiased	Biased
2	Tyr	1.1	0.9	3.4	0.5
9	Tyr	1.2	0.5	2.3	0.8
25	Trp	1.5	0.4	1.6	0.5
33	His	1.4	0.3	4.9	0.7
37	Pro	1.0	1.1	1.1	1.2
40	Tyr	1.1	0.5	2.7	0.9
45	Leu	0.7	0.6	1.7	0.9
53	Pro	1.0	0.4	1.2	0.9
59	Pro	0.7	0.5	1.5	0.6
60	Trp	0.6	0.8	2.7	0.7
62	Phe → Tyr	0	0	0	0
70	Trp	1.5	0.3	7.1	0.9
72	Tyr	0.6	0.3	1.8	1.1
75	Leu	1.9	1.0	5.5	1.2

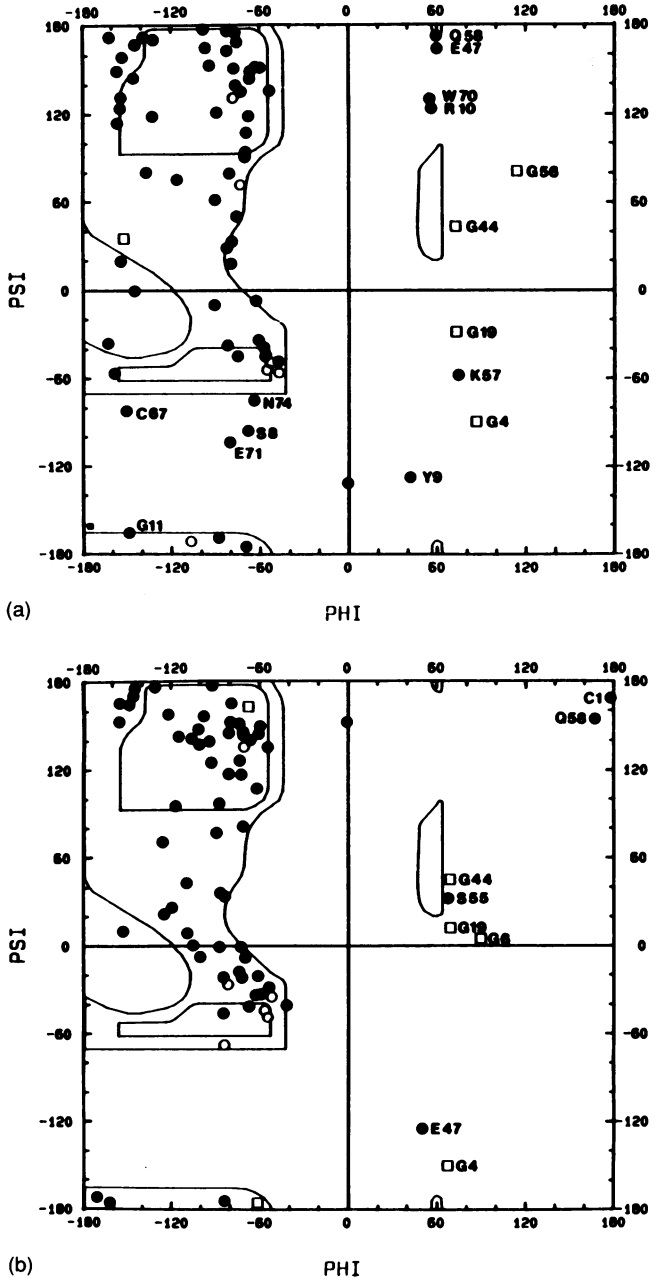

FIGURE 4 Ramachandran plots of the unbiased (*a*) and biased (*b*) models of LPaK9. Glycine residues are boxed and residues with unacceptable angles are identified.

model agree with the torsional angles in the Ramachandran plot. In PGK4, only Met47 falls into nonallowed regions of the Ramachandran plot, whereas in the biased LPaK9 model, Glu47, Ser55, and Cys1 are located in nonallowed regions. In contrast, the Ramachandran plot of the unbiased LPaK9 shows that the dihedral angles are outside of energetically allowed regions (Ser8, Tyr9, Arg10, Glu47, Lys57, Cys67, Trp70, Glu71, Asn74). The structural differences of the unbiased model are also reflected by the positional differences in the main chain of the kringle.

Two possible ligand-binding sites in models of LPaK9

The models predict two different possible ligand-binding sites for LPaK9. In the unbiased LPaK9 model, residues in positions 57, 60, 62, 69, 70, and 72, which are integral to the ligand-binding region of the kringle, may be forced into another energetically favorable conformation. The significant residue changes occurring in the ligand-binding region of LPaK9 are Trp32, Arg35, Ser55, Gly56, and Tyr62. Since Asp56 is not conserved in LPaK9, the two models of this kringle present potential ligand-binding sites that are different from the same region in PGK4. The ligand binding site of the unbiased LPaK9 model shows orientations of Trp60 and Trp70 that are different from the corresponding region in PGK4 structure (Fig. 5, orange). In the model of biased LPaK9 (Fig. 5, blue), the amino acid side chains of Trp60 and Trp70 are positioned as in PGK4. In the LPaK9 models, Trp32 and Arg35, which replace Arg32 and Lys35 of PGK4, respectively, appear to influence the gross fidelity of the PGK4 ligand-binding site by their overall mass, charge characteristics, and orientation. Both models suggest that LPaK9 binds a bulkier and nonpolar ligand such as an aromatic residue. The ligand-binding sites of both LPaK9 models are more spacious than that of PGK4 and may even accommodate two residues. Although the two LPaK9 models were minimized with different amino acids placed in the ligand-binding site space-fill experiments and in docking experiments with fragments of apo B-100, only results of experiments with the biased LPaK9 model, which may be a more valid representation of the actual structure, are discussed.

Modeling of LPaK9 with free amino acid ligands

Expectedly, phenylalanine, proline, and leucine were found to "fit" well as free ligands in the trough formed

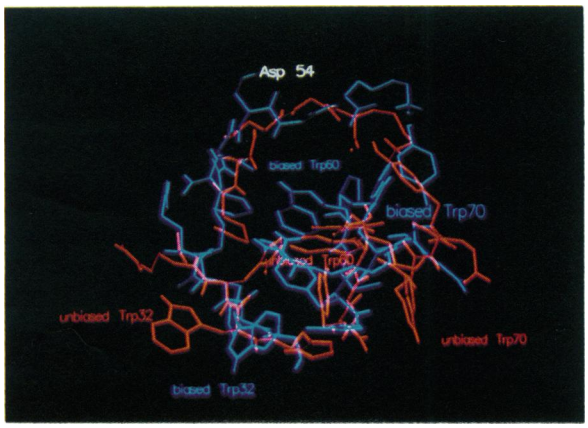

FIGURE 5 Stick models of the binding regions of the unbiased (*orange*) and biased (*blue*) models of LPaK9. A striking difference is evident in the putative ligand-binding regions of the models that reflect the rms values for Trp 60, Trp 70, and Tyr 72, shown in Table 3.

by Trp60 and Trp70 of the ligand-binding site. Free ligands of lysine and arginine were observed to be rotated as much as 90° when placed with the carboxyl moieties proximal to the cationic region of the binding site. In some cases the guanidino group of arginine and the ϵ-amino group of lysine were observed to fold back to form a salt bridge with the delta-carboxyl group of Asp54. In these instances, a major portion of the ligand-binding site is left unoccupied in the nonpolar region (i.e., Gly56 and Trp60 and Trp70 near Tyr72).

Modeling of possible interactions between LPaK9 and apo B-100

In LPaK9, Cys67, which probably forms the disulfide with apo B-100, is on the surface of the kringle between the putative ligand-binding site and a region rich in aromatic and hydrophobic residues. The latter region is composed of Met28, Ile29, Pro30, Trp32, and Pro66. A multi-faceted interaction between LPaK9 and apo B-100 that involves the ligand-binding site, the kringle surface adjacent to the binding site and Cys67, and the disulfide bond appeared to be possible. The polypeptide fragments of apo B-100 shown in Table 1 were subjected to energy minimization with the LPaK9 model as a docking template to determine which fragment(s) had the best fit with the ligand binding site and surface of LPaK9.

Assessment of "best fit" was based on total number of contacts between the fragment and the LPaK9 model, lowest potential energy values after minimization, and dihedral main chain angles for the fragment. Apo B-100 fragment P_{3732}SCKLDFR was repeatedly shown to be the most complementary peptide with the kringle. Typically, lower potential energy values were achieved when the kringle with the apo B-100 fragment was subjected to minimization. Although these initial modeling experiments with the shorter fragments were not rigorously minimized, this fragment was shown to have a high number of contacts with Phe3738 having 14 good contacts with Trp60 and Trp70. The high contact total was then confirmed through a second method in docking and minimization experiments in which distance constraints between the cysteine residues were imposed and the kringle was held fixed. Results shown in Table 4 indicate that apo B-100 fragment P_{3732}SCKLDFR had the lowest potential energy and 90 atom contacts, with phenylalanine having 21 with the critical ligand-binding site residues of Trp60 and Trp70. Apo B-100 fragment L_{1393}SCDGSLR displayed 96 atom contacts and the second best potential energy value. The leucine and arginine residues of this fragment had 41 contacts with residues of the ligand-binding site but none with Trp60; arginine showed 33 contacts with kringle residues Gln30, Arg35, Asp54, Ser55, and Gly56 on the periphery of the ligand-binding site. Although this apo B-100 fragment shows many contacts with the kringle model ligand-binding site, practically all contacts are in one region of the site, leaving a large portion of the site vacant. Apo B-100 fragment L_{1083}SCDTKEE also has 95 atom contacts with a good potential energy value; however, the cysteine residue of the fragment is located at an unfavorable distance from the kringle Cys67. Three fragment residues are located within the kringle ligand-binding site. Although many contacts were observed for Tyr1474 for the apo B-100 fragment containing Cys1478, few contacts were evident for the other residues of this fragment. In addition, Met1080, located on the amino terminal side of Cys1085, also displayed a high number of uniform contacts with the kringle ligand-binding site residues, but less complementarity was evident with the other residues in this apo B-100 fragment.

Two approaches were used in modeling the kringle with larger apo B-100 fragments. The best results were obtained in docking experiments and energy minimization of apo B-100 fragment P_{3732}SCKLDFREIQIYKK with the biased LPaK9 model. Results in Table 5 show a uniform interaction between residues of apo B-100 fragment P_{3732}SCKLDFREIQIYKK and the biased LPaK9 model ligand-binding site and surface. About 188 molecular contacts were identified in docking experiments in which the apo B-100 fragment P_{3732}SCKLDFREIQIYKK was contoured on biased LPaK9 model (Figs. 6 and 7). Only 30 contacts were evident between the unbiased LPaK9 and this fragment of apo B-100. Significantly, 139 nonhydrogen bond contacts between the LPaK9 model and this apo B-100 fragment were between 3.5 and 4.0 Å. Forty-nine contacts were <3.5 Å, but there were no van der Waals contacts < 2.9 Å. The first residue of the fragment, Pro3732, was shown to have 26 total contacts with Pro30, His31, Asp65, Pro66, Cys67, and Arg69 of LPaK9. Apo B-100 Pro3732 has 16 nonhydrogen bond contacts between 3.5 and 4.0 Å and 10 contacts < 3.5 Å with the biased model of LPaK9. The ligand residue, Phe3738, displayed 29 total contacts, between 3.5 and 4.0 Å, with only one contact of 2.94 Å at Phe CE2 and Asp54 OD1. This suggests that Phe3738 fits well within the ligand-binding site of this LPaK9 model when it is located at position 7 in this apo B-100 fragment. Contact points shown in Table 5 suggest that LPaK9 Pro30, His31, Pro66, Cys67, and Arg69 may interact with residues in positions 1–4 of the apo B-100 fragment with cysteine in position 3. LPaK9 Gln34, Arg35, Asp54, Gly56, Tyr62, Trp60, and Trp70 form the ligand-binding site and interact with hydrophobic residues such as phenylalanine and leucine in fragment position 7 and with residues in fragment positions 8 and 9. To identify LPaK9 residues that might interact with fragment residues in positions 9–15, dipeptides were added to the fragment residue nearest the cationic periphery of the ligand-binding pocket, i.e., His33, Gln34, and Arg35. Obviously, one possibility was that the fragment polypeptide would be minimized orthogonal or away from the kringle surface. The second possibility was that the fragment polypeptide would build on the

TABLE 4 Apo B-100 cysteine-containing fragments minimized with the biased LPaK9 model

K	I	T	E	V	A	L	M	G	H	L	S	C	D	T
Cys1085							R35	Y62	W32		R69	P66	P30	
TC 69							D54		H33			C67	H31	
TE −2794							G56		Q34				P66	
							W60		R69					
							W70							
							Y72							
						Contacts	21	01	17	00	11	04	19	00
L	S	C	D	T	K	E	E	R	K	I	K	G	V	I
D65	C67		C67	R69	D54	R35	Q34	**Cys1085**						
C67			R69	W70	W60	D54	R35	**TC 95**						
					W70		D54	**TE −2697**						
							Y62							
							R69							
03	04	00	10	07	24	02	45	Contacts						
T	T	Y	D	H	K	N	T	F	T	L	S	C	D	G
Cys1395							Q34	D54	R69	R69	P66	P66	P66	H31
TC 66								W60			C67	C67		W32
TE −2749								Y62						R69
								W70						
						Contacts	04	14	02	02	10	11	01	22
L	S	C	D	G	S	L	R	H	K	F	L	D	S	N
P66		P66	R69	W32	Q34	R35	Q34	**Cys1395**						
C67		C67		H33	R35	D54	R35	**TC 96**						
		R69		R69	Y62	Y62	D54	**TE −2878**						
					R69	W70	S55							
							G56							
09	00	14	09	08	15	08	33	Contacts						
S	S	F	Y	A	K	G	T	Y	G	L	S	C	Q	R
Cys1478							D54	R35	R69		R69	P66	P30	P66
TC 81								S55	D54					
TE −2826								G56	G56					
								K57						
								W60						
								W70						
								Y72						
						Contacts	16	38	05	00	13	02	06	01
L	S	C	Q	R	D	P	N	T	G	R	L	N	G	E
	P66	P66	R69	V68	R69		Q34	**Cys1478**						
		C67		W70	W70		R35	**TC 74**						
							Y62	**TE −2858**						
							R69							
00	01	07	13	12	16	00	25	Contacts						
D	G	I	S	T	S	A	T	T	N	L	K	C	S	L
Cys1635							Q34	R35	W70	R69	R69	P66		
TC 38							R35	D54		W70		C67		
TE −2767							D54	W60						
							Y62	Y62						
								W70						
						Contacts	10	13	02	08	02	03	00	00
L	K	C	S	L	L	V	L	E	N	E	L	N	A	E
	P66	P66	P66	R69	W70	Y62	Q34	**Cys1635**						
		C67	C67	W70		R69	R35	**TC 41**						
			R69			W70		**TE −2787**						
00	03	03	07	05	03	08	12	Contacts						

TABLE 4 (continued)

F	H	V	P	F	T	D	L	Q	V	P	S	C	K	L
Cys3734								D54	R69	C67	P66		P30	
TC 38								W60			C67		P66	
TE −2799								Y62			R69			
								W70						
						Contacts	00	16	03	01	12	00	06	00
P	S	C	K	L	D	F	R	E	I	Q	I	Y	K	K
P30	P30	P66		C67	W70	D54	R35	**Cys3734**						
	H31	C67		V68		G56	D54	**TC 90**						
	P66	R69		R69		W60		**TE −2917**						
	R69			W70		W70								
08	12	09	00	14	04	21	22	Contacts						
K	A	D	Y	V	E	T	V	L	D	S	T	C	S	S
Cys3890							Q34	W70	Y62	R69	P66	P30	P30	I29
TC 69							R35		R69		C67	H31		P30
TE −2806											R69	P66		W32
						Contacts	04	03	20	03	07	07	04	21
S	T	C	S	S	T	V	Q	F	L	E	Y	E	L	N
	P66	R69	R69	Y62	R35	W60		**Cys3890**						
	C67			R69	D54	W70		**TC 49**						
					W60			**TE −2786**						
					Y62									
					W70									
00	06	03	08	08	17	07	00	Contacts						
P	G	K	P	G	I	Y	T	R	E	E	L	C	T	M
Cys4190							R69	R35	W70	R69	H31	C67	P66	
TC 64								D54			P66		C67	
TE −2909								W60			C67			
								Y62			R69			
								R69						
								W70						
						Contacts	04	26	07	15	07	03	02	
E	L	C	T	M	F	I	R	E	V	G	T	V	L	S
P66		P66	R69	W32	W32	R35	R35	**Cys4190**						
C67				R69		Y62	D54	**TC 46**						
						R69		**TE −2853**						
						W70								
03	00	03	05	08	05	05	17	Contacts						

Apo B-100 fragments of eight amino acids were constructed to determine total contacts (**TC**) with residues adjacent to LPaK Cys67, those residues between Cys67 and the ligand-binding site, and residues with the binding site (e.g., P66 is LPaK9 Pro66). A distance constraint of 2.0 Å was imposed between LPaK Cys67 and the cysteine of the apo B-100 fragment (e.g., **Cys1085**). Fragments were constructed two residues per minimization cycle. Final total energy values are shown as **TE**.

surface of the kringle and suggest which kringle residues interact with the docking polypeptide. Thereby, a complement polypeptide sequence would evolve. LPaK9 Met28, Trp32, Gln34, Thr36, and Glu38 would appear to interact with nonpolar residues in fragment positions 11–15. However, a polar residue such as aspartate in fragment position 13 may turn the fragment away from the kringle surface. This result was observed using the apo B-100 fragment containing Cys3890 (Table 5).

Other regions of apo B-100 that contain unbound cysteine residues were also studied, and less atom contact points with less favorable potential energy values were observed after minimization with this LPaK9 model. For example, the graphic in Fig. 8 shows the biased LPaK9 model with apo B-100 fragment $L_{1476}SCQRDPNT_{1484}$ minimized to a location outside the binding site and apart from the kringle. This apo B-100 fragment contains a Pro(1482) that is located

TABLE 5 Apo B-100 Cys3734 and Cys3890 containing fragments minimized with the biased LPaK9 model

I. The apo B-100 15 residue fragment was constructed and contoured on kringle with Phe3738 in ligand-binding site.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	S	C	K	L	D	F	R	E	I	Q	I	Y	K	K
P30	C67	P66	C67	V68	V68	R35	R69	H33	Q34	W32	S27	M28	M28	M28
H31		C67	V68	W70	R69	D54		Q34	R69	H33	M28	T36		
D65			R69		W70	W60		R35		Q34	H31	E38		
P66						Y62		Y62			W32			
C67						W70		R69			H33			
R69											T36			
26	01	05	06	05	26	29	01	39	07	09	16	16	01	01
(05)	(01)			(01)	(03)			(09)	(02)	(03)	(01)			

Cys3734
TC 188
TE −3458

II. The apo B-100 15 residue fragment was constructed by adding dipeptides in the carboxyl-terminal direction to Arg3739 in the kringle ligand-binding site. Each addition step was followed by a minimization cycle.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	S	C	K	L	D	F	R	E	I	Q	I	Y	K	K
P30	P30	P66			R69	G56	R35	Q34		W32	W32	M28		S26
H31	H31	C67			W70	W60	D54	R35		Q34		W32		E38
	P66	R69				W70		R69				T36		
	R69					D54						E38		
09	12	05	00	00	11	15	10	11	00	08	02	13	00	07
(03)	(03)	(02)			(01)		(03)	(05)		(02)		(01)		(03)

Cys3734
TC 103
TE −3351

III. The apo B-100 15 residue fragment was constructed by adding dipeptides in the amino-terminal direction to Val3885 in the kringle ligand-binding site. Each addition step was followed by a minimization cycle.

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
K	A	D	Y	V	E	T	V	L	D	S	T	C	S	S
			Q34	Q34	Q34	Q34	D54	G56	Y62	R69	C67	H31	P30	I29
			N39		R35	R35	G56	K57	R69		R69	P66	H31	P30
			Y40		Y40	D54		W60	W70			C67	D65	W32
						Y62		W70				R69	P66	
								Y72						
00	00	00	06	04	10	18	08	22	25	06	10	12	16	13
			(01)	(02)	(04)	(05)		(01)	(07)	(02)	(01)	(02)	(07)	(02)

Cys3890
TC 150
TE −3127

Contacts between the apo B-100 fragment residues (single letter) and the LPaK9 residues (e.g., Q34) are shown in each column. Hydrogen bond contacts are shown in parenthesis. Total number of contacts (**TC**) between the fragment and LPaK9 include van der Waals and hydrogen bond contacts. The final total energy values (**TE**) shown in the table reflect values after minimization was performed on the LPaK9 and the complete apo B-100 fragment.

four residues from Cys1477. This proline was placed in the hydrophobic ligand-binding site of the kringle, and the disulfide bond between LPaK9 Cys67 and apo B-100 Cys1477 was formed before energy minimization. An unusually high potential energy value was computed before minimization. The minimization process moved the fragment away from the kringle but was constrained by the disulfide bond. Incompatibility between this apo B-100 fragment and the LPaK9 models is suggested by the few contact points with the ligand-binding site residues Trp60 and Trp70. This example also shows the importance of the dihedral angles for residues placed in the ligand-binding site. This program is designed to adjust angles and bond lengths to reach the most favorable energy levels, and, in this instance, the lowest energy levels for both the proline of the apo B-100 fragment and the

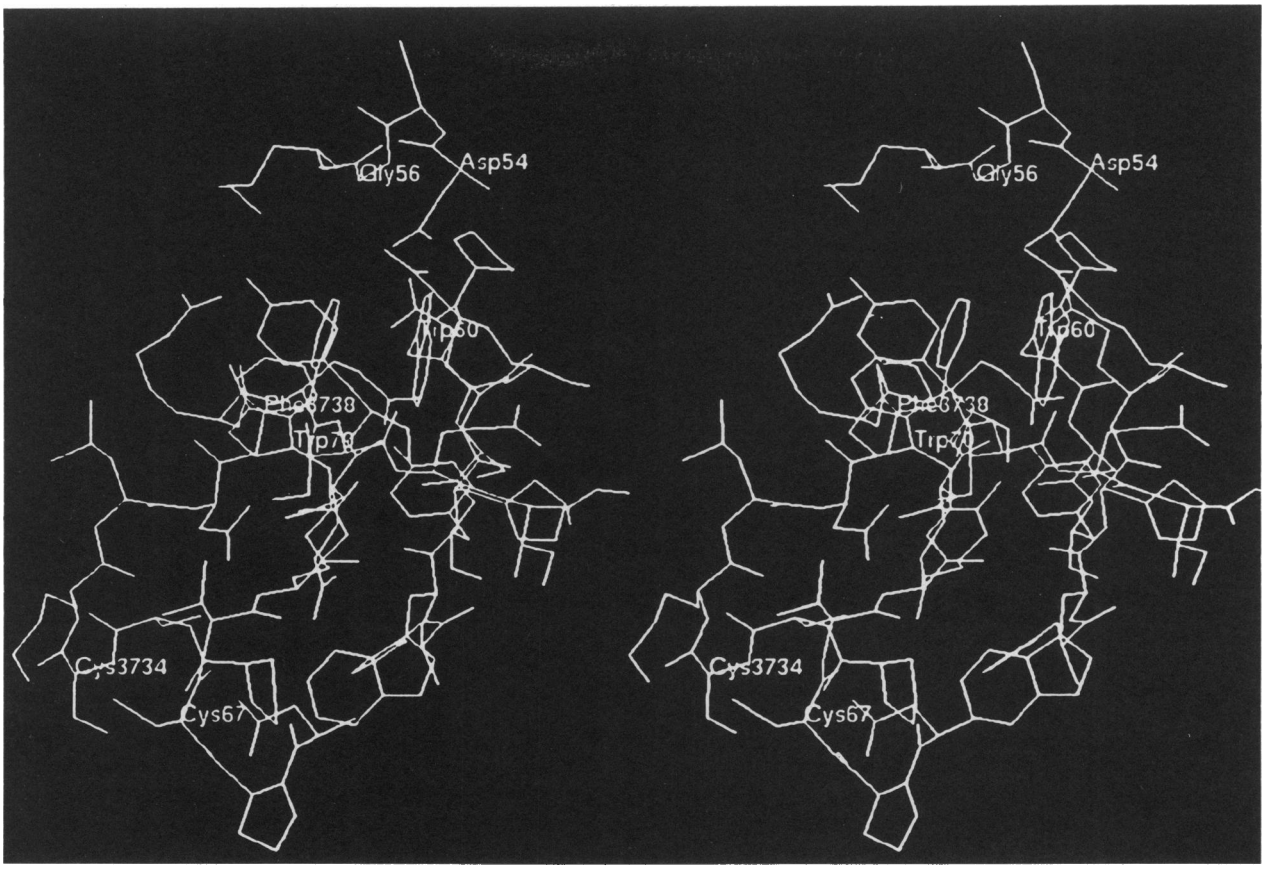

FIGURE 6 Stereogram of the biased LPaK9 model showing residues that interact with apo B-100 fragment P_{3732}SCKLDFREIQIYKK. The fragment was contoured on the kringle with Phe stacked between Trp60 and Trp70 before minimization.

kringle model are achieved only with the fragment away from the kringle.

DISCUSSION

The amino acid sequence of apo[a] inferred by the cDNA sequence (1) suggests the presence of 11 different types of kringles, with 10 kringle types being similar but not identical to PGK4 (17, 40). We have assumed that (*a*) there exists a high degree of homology among all kringles, (*b*) there is the conservation of specific residues in determined positions that define the kringle motif, and (*c*) there is especially high sequence homology between apo[a] kringle sequences and PGK4. Based on these assumptions, it is reasonable to predict that the apo[a] PGK4-like sequences also exist in a kringle conformation. Although there is no direct evidence at present demonstrating that these repeated sequences exist in the typical tri-loop kringle conformation, it is known that prothrombin kringle 1, plasminogen kringles 1 and 4, and tissue plasminogen activator kringle 2, which share ~50% sequence homology, occur in the same kringle conformation. Therefore, operating on this assumption we have developed molecular models of the

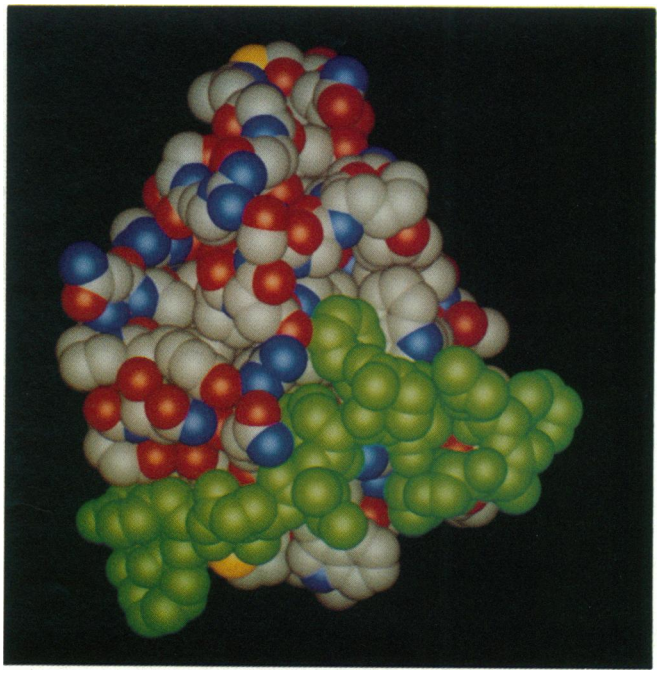

FIGURE 7 Space-filled model of apo B-100 fragment 3732–3745 docked to biased LPaK9.

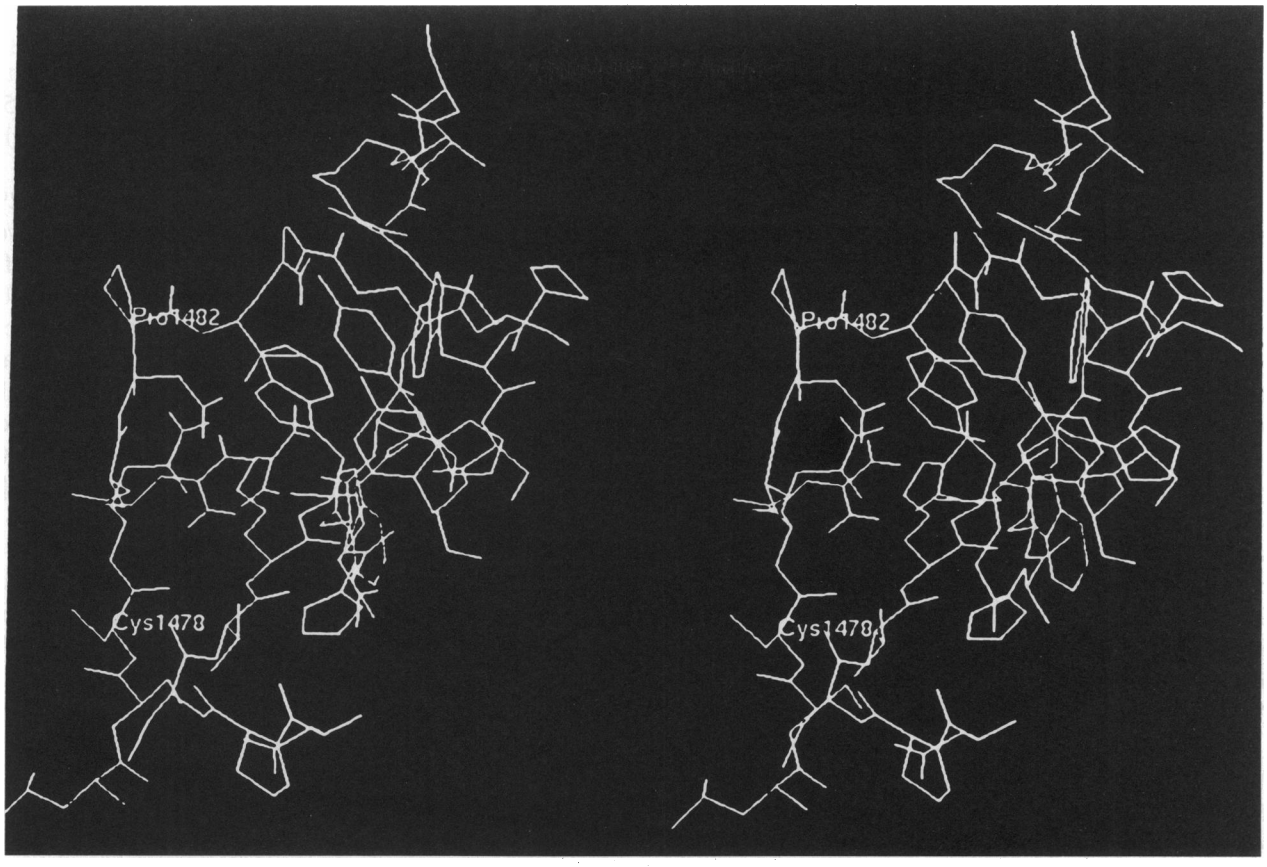

FIGURE 8 Stereogram of the biased LPaK9 with apo B-100 fragment 1476–1489. This apo B-100 fragment contains Pro1481 that is located four residues from Cys1477 and was docked in the hydrophobic ligand-binding site of the kringle. The Pro1481 can be seen outside the binding site with no apparent contact points after energy minimization.

LPaK9 sequence as a kringle to gain a clearer understanding of this sequence both as a kringle type and as an important site on apo[a] for the kringle interaction with apo B-100 in the forming Lp[a]. Our second assumption is that because the overall kringle side chain structure appears to be the same for conserved residues in PGK4 and prothrombin kringle 1 (27), these structures also may be conserved in apo[a]. Although refinement of molecular structures by nonhydrated, static energy minimization can yield secondary and tertiary structures that are energetically favorable but are different from the true structure, at present it is still the best method for evaluating structures that have not been elucidated by x-ray crystallographic or multidimensional nuclear magnetic resonance methods. Furthermore, molecular modeling provides a reasonable prediction of other structures that may actually exist. Our models are based on coordinates of the crystallographic structure for PGK4 and are presented here as predicted structures of LPaK9.

We have used these predicted LPaK9 structures in computer graphics ligand site space-filling and docking experiments with fragments of apo B-100 to visualize possible molecular interactions that may exist between apo B-100 and apo[a]. Several schemes can be considered to explain how apo[a] and apo B-100 interact in the intact Lp[a] particle. The most basic is limited simply to the formation of a single connecting covalent disulfide bond and proposes that the remainder of the apo[a] molecule floats freely in the aqueous environment about the lipoprotein particle. A second mechanism involves, in addition to the connecting disulfide, ligand-binding of the two molecules through the kringle domain of apo[a] and specific binding sites on apo B-100. For example, a lysine-containing region on apo B-100 might bind through ionic and van der Waals forces directly with a ligand-binding site on one or several apo[a] kringles. A third mechanism involves electrostatic interactions between the highly sialylated, negatively charged apo[a] and the positively charged apo B-100 of the particle.

Although it has not yet been demonstrated directly by amino acid sequence analysis, there is strong evidence that apo[a] is covalently linked to apo B-100 (7–11, 14, 37). We have performed experiments that identify indirectly a cysteine residue on apo B-100 that forms a disulfide bond with apo[a]. Our results show that in labeling Lp[a] with F5M, only the Lp[a−] particle was labeled. This suggests that a cysteine residue in LPaK9 thought not to be paired within the apo[a] molecule may form a

disulfide bond with apo B-100 in the intact Lp[a]. Although our results indicate that Cys3734 is labeled in LDL but is not labeled in Lp[a] and suggest that this cysteine residue is involved in disulfide linkage between the two apoproteins, direct evidence from sequence analysis has not yet been obtained. Whether Cys3734 of apo B-100 forms a disulfide bond exclusively with apo[a] has not been determined. It is possible that this cysteine residue of apo B-100 may form a disulfide bond with other apolipoproteins such as apoE or apoD or form a thioester linkage with fatty acids (38).

Several noteworthy aspects of LPaK9 that are evident in its primary structure are then graphically revealed by our molecular models. LPaK9 has a ligand-binding site similar to, yet different from, that of PGK4 and hence may not be involved in binding to a positively charged ligand such as lysine, arginine, or histidine located either on apo B-100 or another protein. Although the majority of the apo[a] kringle types are similar to PGK4, only LPaK10 appears to retain the critical amino acid sequence required for binding a lysine ligand. We have modeled several free amino acids in the binding site of LPaK9 to identify those that best fill and complement the site. Results of space-fill and minimization experiments with single amino acids as free ligands were inconclusive. Although energetically favorable models were attained with some free amino acids, the orientation of amino acids such as proline, lysine, and arginine within the site was visibly not convincing. For example, proline minimizes well with LPaK9 models but only partially fills the apparent ligand-binding site space. Lysine and arginine are minimized into a "curled" conformation to form a salt bridge with Asp54 of the kringle and hence only a part of the site is occupied. Space-fill and minimization experiments with amino acids leucine and phenylalanine, as free ligands, in the ligand-binding site of the biased LPaK9 model showed energetically favorable and visibly acceptable models; low potential energy values for the complex with many atom contacts between the ligand and amino acids that form the ligand-binding site. However, it is unlikely that this type of interaction is of physiologic importance. The ligand amino acid is more likely part of a polypeptide and involves other adjacent amino acids.

In Lp[a], we considered a mechanism that involves both covalent disulfide bonding and noncovalent interactions between a specific site on apo B-100 and a specific amino acid such as leucine, methionine, tyrosine, or phenylalanine. Although 14 apo B-100 fragments (Table 1) were examined, only six regions were expected to be compatible with LPaK9: Cys1395 with Phe1391 or Leu1399, Cys1478 with Tyr1474, Cys1635 with Val1639, Cys3734 with Phe3738, Cys3890 with Leu3886 or Val3894, and Cys4190 with Ile4194. Two apo B-100 fragments containing Cys3734 and Cys3890 are part of an epitope recognized by monoclonal antibodies, Bsol15 and Bsol20, reported by Zawadzki et al. (39). Their results show a slight increase in immunoreactivity for these antibodies with the apo[a]-free Lp[a−] as compared with the intact Lp[a]. The location of this epitope on apo B-100 is at 3506–4082 and contains 576 amino acids. Only 29 residues, ~5%, contained within this apo B-100 fragment were not released into an aqueous medium after trypsin digestion of LDL (29). Furthermore, ~35% of the remaining residues were found in both hydrophilic and hydrophobic media after trypsin digestion of LDL (29), which suggests a flexibility in conformation as well as in its microenvironment. Although it appears that all the disulfide bonds of apo B-100 occur on the surface in the trypsin-releaseable peptides of the LDL particle (34), there are no disulfides located in this region of the apo B-100 molecule. It is possible that in Lp[a], the interaction between apo[a] and apo B-100 may be restricted to LPaK9 and to 15 or less residues on apo B-100 within this epitope.

Systematic energy minimizations of free ligands and ligands as part of polypeptides filling the ligand-binding space in the biased LPaK9 model strongly suggest that fragment P_{3732}SCKLDFREIQIYKK is the site of interaction between apo[a] and apo B-100 in Lp[a]. This interaction is based on complementarity that involves (*a*) the covalent disulfide bond formed between apo B-100 Cys3734 and LPaK9 Cys67, (*b*) the interaction of apo B-100 Phe3738 with Trp60 and Trp70 within the ligand binding site of LPaK9, and (*c*) the electrostatic and van der Waals interactions between the apo B-100 fragment and the surface of LPaK9 adjacent to the ligand-binding site and Cys67. Molecular energy refinement (dihedral energy, bond length energy, van der Waal energy) was enhanced by the presence of the apo B-100 fragment, and many contact points between these LPaK9 models and this apo B-100 fragment were identified. Although minimization was not hindered by the combination of the LPaK9 model with a noncompatible apo B-100 fragment, the fragment and kringle typically moved apart and basically minimized separately. Molecular energy refinement of noncomplementary structures was therefore accomplished with the two molecules apart. An example of this result is demonstrated with fragment L_{1476}SCQRDPNTGRLNGE (Fig. 8).

Proline, an amino acid that is thought to be more important than lysine in the binding of apo[a] and Lp[a] to biotinylated-LDL (32), was also modeled with LPaK9, alone and within apo B-100 fragment 1476–1490. In our models, proline appears to occupy only a fraction of the apparent space in the binding site of LPaK9. As a part of apo B-100 fragment 1476–1490, proline appeared to be minimally compatible with LPaK9. It is likely that proline, when added in excess, interferes with the interactions between apo[a] and biotinylated-LDL by competing with a proline on apo B-100 that binds to a different apo[a] kringle type such as the highly repeated type 2 (LPaK2). The critical anionic residue, Asp56 in PGK4, is replaced by glycine in

LPaK2, a change that influences the dipolarity and size of the kringle binding site. There are 78 proline residues in the trypsin-releasable peptides of apo B-100 from LDL (34), 56% of which are located near the NH_2-terminal end. Little is known about the function of proline residues in apo B-100. It is possible that some prolines are located in exposed structural loci such as beta-turns that would render these residues sterically accessible for interacting with ligand-binding sites of apo[a] kringles. Alternatively, it is possible that excess proline disrupts the hydrophobic core of the kringle (27) formed by the numerous proline, histidine, and aromatic residues that may be conserved in some apo[a] kringles and, thereby, alter the geometry of the kringle ligand-binding site. In bovine prothrombin kringle 1 and PGK4, there are histidine and proline residues that form discrete clusters (27). Exogenous proline or hydroxyproline may disrupt these histidine/proline clusters. If such clusters also exist in apo[a] kringles, altering their structure by proline would be expected to interfere with their binding capacity. One such cluster includes residues Pro30, His31, and Pro66, which are conserved in apo[a] kringle types 1 through 10. In docking experiments with the biased LPaK9 model and apo B-100 fragment P_{3732}SCKLDFREIQIYKK, Pro3732 appears to have contact points with Pro30, His31, Asp65, Pro66, Cys67, and Arg69. There are several other key residue alterations in this region of the kringle such as Ile29 and Trp32 that along with Met28, Pro66, and Val68 create a more hydrophobic environment. This region may interact with the apolar regions adjacent to the Cys3734 apo B-100 fragment and/or the lipid phase of the LDL particle. Alternatively, this site may be a ligand-binding region in some kringles.

CONCLUSIONS

Based on energy minimization calculations, it is clear that an LPaK9 sequence is capable of assuming a molecular conformation similar to that of PGK4 determined crystallographically. Two molecular models of LPaK9 were obtained through minimization methods with energetically favorable values that are comparable with those of PGK4. The more plausible and acceptable model of LPaK9, with few unfavorable dihedral angles, was the biased one generated by simply altering the 16 amino acid substitutions on the complete PGK4 molecule. Major sites on the kringle such as the ligand binding region are influenced by the character of the residues located at or near that region. Therefore, it is not surprising that, since PGK4 and LPaK9 have different amino acid residues in this region, the two kringles have binding sites of apparently different ligand specificity and/or affinity.

Molecular docking experiments with models of LPaK9 and apo B-100 fragments that contain a free cysteine lead us to propose that the interaction between apo[a] and apo B-100 in Lp[a] involves three aspects: (*a*) LPaK9 forms a covalent disulfide bond at Cys3734; (*b*) there are molecular contacts, electrostatic interactions, hydrogen bonds, and van der Waals interactions between specific amino acids on the face of LPaK9 and that of LDL apo B-100; and (*c*) a kringle-ligand complex forms between apo B-100 Phe3738 and the ligand-binding site of LPaK9.

Supported in part by National Institutes of Health (NIH) grant HL-32971 to J. D. Morrisett, NIH grant 2S07RR-05425-29 to J. Guevara, Jr., HL-25942 to A. Tulinsky, GM-27061 to B. V. V. Prasad, and The W. M. Keck Center for Computational Biology. Some molecular modelling was performed using facilities of the Molecular Biology Information Resource, Department of Cell Biology, and 3DEM Resource Center (supported by NIH grant RR-02250 and The W. M. Keck Center Foundation), Baylor College of Medicine.

Received for publication 18 June 1992 and in final form 25 September 1992.

REFERENCES

1. McLean, J. W., J. E. Tomlinson, W.-J. Kuang, D. L. Eaton, E. Y. Chen, G. M. Fless, A. M. Scanu, and R. M. Lawn. 1987. cDNA sequence of human apolipoprotein(a) is homologous to plasminogen. *Nature (Lond.)*. 330:132–137.
2. Dahlen, G. H., J. R. Guyton, M. Attar, J. A. Farmer, J. A. Kautz, and A. M. Gotto, Jr. 1986. Association of levels of lipoprotein Lp[a], plasma lipids, and other lipoproteins with coronary artery disease documented by angiography. *Circulation*. 74:758–765.
3. Hoefler, G., F. Harnoncourt, E. Paschke, W. Mirtl, K. H. Pfeiffer, and G. M. Kostner. 1988. Lipoprotein Lp(a): a risk factor for myocardial infarction. *Arteriosclerosis*. 8:398–401.
4. Zenker, G., P. Koltringer, G. Bone, K. Niederkorn, K. Pfeiffer, and G. Jurgens. 1986. Lipoprotein(a) as a strong indicator for cerebrovascular disease. *Stroke*. 17:942–945.
5. Murai, A., T. Miyahara, N. Fujimoto, M. Matsuda, and M. Kameyama. 1986. Lp(a) lipoprotein as a risk factor for coronary heart disease and cerebral infarction. *Atherosclerosis*. 59:199–204.
6. Ehnholm, C., H. Garoff, O. Renkonen, and K. Simons. 1972. Protein and carbohydrate composition of lp(a) lipoprotein from human plasma. *Biochemistry*. 11:3229–3232.
7. Gaubatz, J. W., C. Heideman, A. M. Gotto, Jr., J. D. Morrisett, and G. H. Dahlen. 1983. Human plasma lipoprotein [a] structural properties. *J. Biol. Chem.* 258:4582–4589.
8. Gaubatz, J. W., M. V. Chari, M. L. Nava, J. R. Guyton, and J. D. Morrisett. 1987. Isolation and characterization of the two major apoproteins in human lipoprotein[a]. *J. Lipid Res.* 28:69–79.
9. Armstrong, V. W., A. K. Walli, and D. Seidel. 1985. Isolation, characterization, and uptake in human fibroblasts of an apo(a)-free lipoprotein obtained on reduction of lipoprotein(a). *J. Lipid Res.* 26:1314–1323.
10. Fless, G. M., M. E. ZumMallen, and A. M. Scanu. 1986. Physicochemical properties of apolipoprotein(a) and lipoprotein(a−) derived from the dissociation of human plasma lipoprotein(a). *J. Biol. Chem.* 261:8712–8718.
11. Seman, L. J., and W. C. Breckenridge. 1986. Isolation and partial characterization of apolipoprotein(a) from human lipoprotein (a). *Biochem. Cell Biol.* 64:999–1009.

12. Eaton, D. L., G. M. Fless, W. J. Kohr, J. W. McLean, Q-T. Xu, C. G. Miller, R. M. Lawn, and A. M. Scanu. 1987. Partial amino acid sequence of apolipoprotein(a) shows that it is homologous to plasminogen. Proc. Natl. Acad. Sci. USA. 84:3224–3228.

13. Kratzin, H., V. W. Armstrong, M. Niehaus, N. Hilschmann, and D. Seidel. 1987. Structural relationship of an apolipoprotein (a) phenotype (570 kDa) to plasminogen: homologous kringle domains are linked by carbohydrate-rich regions. *Biol. Chem. Hoppe-Seyler* 368:1533–1544.

14. Utermann, G., and W. Weber. 1983. Protein composition of Lp[a] lipoprotein from human plasma. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 154(2):357–361.

15. Cushing, G. L., J. W. Gaubatz, M. L. Nava, B. J. Burdick, T. M. A. Bocan, J. R. Guyton, D. Weilbaecher, M. E. DeBakey, G. M. Lawrie, and J. D. Morrisett. 1989. Quantitation and localization of apolipoproteins [a] and B in coronary artery bypass vein grafts resected at re-operation. *Arteriosclerosis.* 9:593–603.

16. Rath, M., A. Niendorf, T. Reblin, M. Dietel, H.-J. Krebber, and U. Beisiegel. 1989. Detection and quantification of lipoprotein(a) in the arterial wall of 107 coronary bypass patients. *Arteriosclerosis.* 9:579–592.

17. Morrisett, J. D., J. W. Gaubatz, R. D. Knapp, and J. G. Guevara, Jr. 1990. Structural properties of apo(a): a major apoprotein of human lipoprotein(a). *In* Lipoprotein(a). A. Scanu, editor. Academic Press, Inc., San Diego. 53–74.

18. Magnusson, S., L. Sottrup-Jensen, T. E. Petersen, and H. Claeys. 1975. The primary structure of prothrombin, the role of vitamin K in blood coagulation, and thrombin-catalyzed negative feedback for limiting the activation of prothrombin. *In* Prothrombin and Related Coagulation Factors (Boerhaave Symposium, May 1974). H. C. Hemker and J. Veltkamp, editors. University Press, Leiden. 25.

19. Magnusson, S., L. Sottrup-Jensen, T. E. Petersen, G. Dudek-Wojciechowska, and H. Claeys. 1976. Homologous "kringle" structures common to plasminogen and prothrombin. Substrate specificity of enzymes activating prothrombin and plasminogen. *In* Proteolysis and Physiological Regulation. D. W. Ribbons and K. Brew, editors. Academic Press, New York. 203–238.

20. Furie, B., and B. C. Furie. 1988. The molecular basis of blood coagulation. *Cell.* 53:505–518.

21. Hedner, V., and E. W. Davie. 1989. Introduction to hemostasis and the vitamin K-dependent coagulation factors. Part 14. Blood and blood-forming tissue. *In* The Metabolic Basis of Inherited Disease. C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, editors. McGraw-Hill Info. Serv. Co., New York. 2107–2134.

22. Park, C. H., and A. Tulinsky. 1986. Three-dimensional structure of the kringle sequence: structure of prothrombin fragment 1. *Biochemistry.* 25:3977–3982.

23. Tulinsky, A., C. H. Park, B. Mao, and M. Llinas. 1988. Lysine/fibrin binding sites of kringles modeled after the structure of kringle 1 of prothrombin. *Proteins Struct. Funct. Genet.* 3:85–96.

24. Tulinsky, A., C. H. Park, and E. Skrzypczak-Jankun. 1988. Structure of prothrombin fragment 1 refined at 2.8 A resolution. *J. Mol. Biol.* 202:885–901.

25. Mulichak, A. M., and A. Tulinsky. 1990. Structure of the lysine-fibrin binding site of human plasminogen kringle 4. *Blood Coag. Fibrinol.* 1:673–679.

26. Seshadri, T. P., A. Tulinsky, E. Skrzypczak-Jankun, and C. H. Park. 1991. Structure of bovine prothrombin fragment 1 refined at 2.25 A resolution. *J. Mol. Biol.* 220:481–494.

27. Mulichak, A. M., A. Tulinsky, and K. G. Ravichandran. 1991. Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 A resolution. *Biochemistry.* 30:10576–10588.

28. Wu, T-P., K. Padmanabhan, A. Tulinsky, and A. M. Mulichak. 1991. The refined structure of the ϵ-aminocaproic acid complex of human plasminogen kringle 4. *Biochemistry.* 30:10589–10594.

29. de Vos, A. M., M. H. Ultsch, R. F. Kelley, K. Padmanabhan, A. Tulinsky, M. L. Westbrook, and A. A. Kossiakoff. 1992. Crystal structure of kringle 2 of tissue plasminogen activator at 2.4 resolution. *Biochemistry.* 31:270–279.

30. Ye, S. Q., V. N. Trieu, D. L. Stiers, and W. J. McConathy. 1988. Interactions of low density lipoprotein$_2$ and other apolipoprotein B-containing lipoproteins with lipoprotein(a). *J. Biol. Chem.* 263:6337–6343.

31. Trieu, V. N., and W. J. McConathy. 1990. Lipoprotein(a) binding to other apolipoprotein B containing lipoproteins. *Biochemistry.* 29:5919–5924.

32. Trieu, V. N., T. F. Zioncheck, R. M. Lawn, and W. J. McConathy. 1991. Interaction of apolipoprotein(a) with apolipoprotein B-containing lipoproteins. *J. Biol. Chem.* 266:5480–5485.

33. Yang, C-Y., T. W. Kim, S-a. Weng, B. Lee, M. Yang, and A. M. Gotto, Jr. 1990. Isolation and characterization of sulfhydryl and disulfide peptides of human apolipoprotein B-100. *Proc. Natl. Acad. Sci. USA.* 87:5523–5527.

34. Yang, C-Y., Z-W. Gu, S-a. Weng, T. W. Kim, S-H. Chen, H. J. Pownall, P. M. Sharp, S-W. Liu, W-H. Li, A. M. Gotto, Jr., and L. Chan. 1989. Structure of apolipoprotein B-100 of human low density lipoproteins. *Arteriosclerosis.* 9:96–108.

35. Coleman, R. D., T. W. Kim, A. M. Gotto, Jr., and C-y. Yang. 1990. Determination of cysteine on low-density lipoproteins using the fluorescent probe, 5-iodoacetamidofluorescein. *Biochim. Biophy. Acta.* 1037:129–132.

36. Sack, J. S. 1988. CHAIN—a crystallographic modeling program. *J. Mol. Graphics* 6:224–225.

37. Sommer, A., R. Gorges, G. M. Kostner, F. Paltauf, and A. Hermetter. 1991. Sulfhydryl-selective fluorescence labeling of lipoprotein(a) reveals evidence for one single disulfide linkage between apoproteins(a) and B-100. *Biochemistry.* 30:11245–11249.

38. Huang, G., D. M. Lee, and S. Singh. 1988. Identification of the thiol ester linked lipids in apolipoprotein B. *Biochemistry.* 27:1396–1400.

39. Zawadzki, Z., F. Tercé, L. J. Seman, R. T. Theolis, W. C. Breckenridge, R. W. Milne, and Y. L. Marcel. 1988. The linkage with apolipoprotein(a) in lipoprotein(a) modifies the immunochemical and functional properties of apolipoprotein B. *Biochemistry.* 27:8474–8481.

40. Guevara, J., Jr., R. D. Knapp, S. Honda, R. S. Northup, and J. D. Morrisett. 1992. A structural assessment of the apo[a] protein of human lipoprotein[a]. Proteins. 12:188–199.