Conformational analysis of apolipoprotein A-I and E-3 based on primary sequence and circular dichroism

Robert T. Nolte and David Atkinson

Departments of Biophysics and Biochemistry, Housman Medical Research Center, Boston University School of Medicine, Boston, Massachusetts 02118 USA

ABSTRACT The primary and secondary structure of human plasma apolipoprotein A-I and apolipoprotein E-3 have been analyzed to further our understanding of the secondary and tertiary conformation of these proteins and the structure and function of plasma lipoprotein particles. The methods used to analyze the primary sequence of these proteins used computer programs: (*a*) to identify repeated patterns within these proteins on the basis of conservative substitutions and similarities within the physicochemical properties of each residue; (*b*) for local averaging, hydrophobic moment, and Fourier analysis of the physicochemical properties; and (*c*) for secondary structure prediction of each protein carried out using homology, statistical, and information theory based methods. Circular dichroism was used to study purified lipid-protein complexes of each protein and quantitate the secondary structure in a lipid environment. The data from these analyses were integrated into a single secondary structure prediction to derive a model of each protein. The sequence homology within apolipoproteins A-I, E-3, and A-IV is used to derive a consensus sequence for two 11 amino acid repeating sequences in this family of proteins.

INTRODUCTION

The plasma apolipoproteins are a diverse group of proteins responsible for many functions in lipid metabolism. The apolipoproteins maintain lipoprotein particle structure, act as cofactors for enzymes, and are the ligands for receptors involved in the cellular targeting of lipoprotein particles. Apolipoprotein A-I (apoA-I;[1] 28 kD, 243 amino acids [AA]) (Brewer et al., 1978) is the major protein component (70%) of plasma high density lipoprotein (HDL) particles (Scanu et al., 1969). ApoA-I is a potent activator of the enzyme lecithin cholesterol acyl transferase (LCAT), a key enzyme in the metabolism of cholesterol (Fielding et al., 1972). Apolipoprotein E-3 (apoE-3; 34 kD, 299 AA), the major isoform of apoE (Rall et al., 1982), is synthesized primarily in the liver and is found on chylomicron, remnant, very low density lipoprotein (VLDL), and HDL particles (Mahley et al., 1984). ApoE-3 contains a region with receptor binding activity to the cellular low density lipoprotein (LDL) (B/E) receptor and putative E receptor (Mahley et al., 1981) and therefore plays a critical role in cellular targeting of lipoprotein particles, particularly the clearance of chylomicron remnants. The related glycoprotein, apolipoprotein A-IV (apoA-IV; 46 kD, 376 AA) (Karathanasis et al., 1986), is synthesized predominantly by the intestine and is associated with chylomicrons and HDL. Lower lipid binding affinity and a tendency to self associate causes apoA-IV to exist also in a nonlipid associated pool in the plasma. The exact role of apoA-IV has not been well established; however, it has been shown to activate the enzyme LCAT (Steinmetz and Utermann, 1985). A knowledge of the structure of these proteins and the molecular details of their interactions with lipids is fundamental to an understanding of the formation, function, and stabilization of lipoprotein particles.

The analysis of apolipoprotein primary sequences has provided many conceptual insights into their structures. The primary sequences of apoA-I, apoE-3, and apoA-IV have been shown to contain tandemly repeated patterns of 11/22 amino acids that comprise most of these molecules (McLachlan, 1977; Boguski et al., 1986; Lou et al., 1986). Segrest and co-workers have suggested that these repeats fold into an amphipathic α-helix (Segrest et al., 1974; Segrest, 1977; Segrest and Feldmann, 1977). The assignment of these repeats to helix has been used as the basis for models of apolipoprotein conformation (Segrest et al., 1974; Sparrow and Gotto, 1982; Brasseur et al., 1990). Sequence variations within individual repeats, which may have significant implications for the ability of a region to fold into helix, generally have been ignored in this approach. However, since any amphipathic structure has the potential to bind to the lipid surface of plasma lipoprotein particles, other structures need to be considered. The identification of several classes of repeats based on sequence, which may exhibit different conformational tendencies, may have important consequences for these models.

Statistically based secondary structure prediction methods have been used as a basis for modeling of the apolipoproteins (Andrews et al., 1976; Mahley, 1988; Karathanasis et al., 1986). However, in general, this approach has been limited to the accuracy of a single prediction method, the Chou and Fasman algorithm (Chou

Address correspondence to Dr. Atkinson.

[1] *Abbreviations used in this paper:* AA, amino acid; AHL, amphipathic helical level; ApoA-I, apolipoprotein A-I; ApoA-IV, apolipoprotein A-IV; ApoE-3, apolipoprotein E; ApoLp-III, apolipophorin-III; CD, circular dichroism; CO, cholesterol oleate; DMPC, dimyristoyl phosphatidylcholine; DPPC, dipalmitoyl phosphatidylcholine; FPS, Fourier power spectra; GOR, Garnier, Osguthorpe, and Robson; HDL, high density lipoprotein; LCAT, lecithin cholesterol acyl transferase.

0006-3495/92/11/1221/19 $2.00

and Fasman, 1978), which uses a globular protein derived data set. Other statistically based methods with similar accuracies have not been applied to the apolipoprotein sequences.

The analysis of the hydrophobic repeat frequency to detect potential amphipathic structures is directly applicable to this protein class and may aid in a more accurate prediction of apolipoprotein conformation. One-dimensional Fourier analysis of the hydropathies (Finer-Moore and Stroud, 1984) has been applied to several of the apolipoprotein sequences by De Loof et al. (1987). However, this analysis has not been applied to the determination of a folding model for the apolipoproteins.

The assignment of structures based on local homology to a database of solved structures (Levin et al., 1986) has not been applied to these sequences. Molecular structures of apolipophorin-III (apoLp-III), an insect apolipoprotein (Breiter et al., 1991), and a portion of ApoE-3 (Wilson et al., 1991) have been reported recently at a resolution of 2.5 Å. These studies provide an appropriate database for the assignment of structure in the apolipoproteins.

In this study, an analysis of the inter- and intrasequence homology, which incorporates conservative substitution and background correction, has been developed and applied to the sequences of apoA-I, apoE-3, and apoA-IV. This analysis has been used to refine the location of two 11 amino acid repeats in this group of proteins. A consensus, representing an idealized sequence of each repeat, has been derived and its properties examined. In addition, an integrated approach to the conformational analysis of the apolipoproteins has been developed. The integrated approach incorporates prediction methods based on statistical analysis, information theory, and sequence homology. Analysis of the physicochemical properties and repeated patterns in these properties, together with secondary structure quantitation using circular dichroism of apolipoprotein/lipid complexes, are incorporated. This approach has been used to predict the secondary structure and possible tertiary folding patterns of apoA-I and apoE-3 in a lipid environment.

MATERIALS AND METHODS

Materials

Dimyristoyl phosphatidylcholine (DMPC; Sigma Chemical Co., St. Louis, MO) and 1-^{14}C-dimyristoyl phosphatidylcholine (^{14}C-DMPC; Amersham Corporation, Arlington Heights, IL) were checked for purity by thin-layer chromatography (TLC) in chloroform:methanol:water:acetic acid (65:25:4:1) on precoated Silica Gel H plates (Analtech, Newark, DE). ApoA-I was purified from human plasma HDL as described by Wetterau and Jonas (1982). ApoE-3 was a gift from Dr. Karl Weisgraber (Gladstone Foundation, San Francisco, CA). Each apolipoprotein was shown to migrate as a single band on a 10% polyacrylamide gel by Coomassie blue staining (Shore et al., 1980).

Preparation of lipid-protein recombinants

Lipid-protein complexes were formed by the addition of 3.0 ml sample buffer (0.15% KCl, 0.10% tris(hydroxymethyl)-aminomethane [Tris]-HCl, 0.025% NaN_3) to a desiccated film containing 1425 μg DMPC and ~0.1 μCi of ^{14}C-DMPC. This mixture was vortexed, and 500 μg apoA-I or apoE-3 (in ~0.2 ml sample buffer) was added to each by subsurface injection. The suspension was incubated overnight at 25°C with mild agitation. The resulting complexes were purified on a 2.6 × 40-cm Sepharose CL-4B column (Pharmacia Inc., Piscataway, NJ) and eluted with sample preparation buffer. The peak column fraction was concentrated using YM-10 Diaflo ultrafiltration membranes (Amicon Division of W.R. Grace and Co., Danvers, MA). The purified lipid-protein complexes were characterized by negative staining electron microscopy.

Circular dichroism analysis

The circular dichroic spectra of DMPC/apoA-I and DMPC/apoE-3 complexes in sample preparation buffer were recorded on a Cary model 61 CD spectropolarimeter (Varian, Palo Alto, CA). Three spectra were recorded for each sample in the far ultraviolet from 207 to 250 nm at a temperature of 24°C in a 0.5-cm cell at a protein concentration of ~0.15 mg/ml. Mean molar ellipticity was calculated from three spectra at 1-nm intervals after corrections for baseline. DMPC at the concentrations used in these studies did not contribute significantly to the CD spectrum (Walsh and Atkinson, 1986). Protein concentrations used in the molar ellipticity calculation were obtained from the samples at the conclusion of spectral recording using a modified Lowry procedure (Markwell et al., 1978).

Secondary structure quantitation was carried out using LINEQ, a nonlinear constrained least-squares curve fitting program by Mao and Wallace (1984).

Sequence identity calculations

A program that incorporates an amino acid comparison function was developed to quantitate the sequence homology between two proteins. The comparison function allows conservative substitutions, which may preserve structural properties, to be considered in the matching criteria. These criteria have not been applied to the sequences of the apolipoproteins and have the potential to locate segments that show only moderate amino acid identity yet are structurally similar because of conservative amino acid substitution (Argos et al., 1983).

Monte Carlo calculations were carried out to determine the statistical significance of observed sequence similarities by determining the frequency of residue match above random levels (Dayhoff et al., 1983). This was achieved by determining the average homology of one protein in the comparison with a random sequence with the same amino acid composition of the second protein. The process of randomizing the sequence and calculating the similarities was carried out until the mean similarity sampled from the comparison matrix had converged with a variance of $<1 \times 10^{-4}$. The proteins examined in this study required ~250 iterations to converage at this level.

The final similarity comparisons are reported as a gray-scaled two-dimensional matrix. Each position in the matrix represents the calculated mean homology between the residues in each window of the sequences centered around the indexes that describe that matrix position. The value reported is the similarity reduced at each position by the Monte Carlo determined background at that position.

Physicochemical analysis

The primary sequences of the apolipoproteins were analyzed with a program entitled "Computational Analysis of Physicochemical Property Scales" or CAPPS. The program was developed to calculate local windowed averages of physical properties (Kyte and Doolittle, 1982),

average moment calculations (Eisenberg et al., 1982; Eisenberg et al., 1984), the calculation of normalized 1D-Fourier transformations (power spectra) (Finer-Moore and Stroud, 1984), and the calculation of the interface level of a helix based on the amino acid distribution. Each of these analyses could be performed on any scale of amino acid properties. In the current study, bulk (Jones, 1975), conformational probabilities (Chou and Fasman, 1978), and hydrophobicity (Hopp and Woods, 1981; Kyte and Doolittle, 1982; Eisenberg, 1984; Cornette et al., 1987) were used. The windowed averaging calculation allowed for variable window size with separate specification for the number of residues included on the left and right sides of the "reporter" residue. The window size was reduced automatically on the left side at the NH_2-terminus and on the right side at the COOH-terminus to minimize edge effects at the ends of the sequence.

Moment calculations are reported as the average vector sum with the individual vector magnitudes derived from the selected amino acid scale and the relative orientation determined by residue position as follows:

$$\mu_{H_i} = \left\{ \left[\sum_{n=i-w_l}^{i+w_r} H_n \cos(n \times X) \right]^2 + \left[\sum_{n=i-w_l}^{i+w_r} H_n \sin(n \times X) \right]^2 \right\}^{1/2}$$

where μ_{H_i} is the moment at the ith residue, X is the angle of repeat (100° for α-helix), i is the sequence number of the current residue, w_r and w_l are the right and left window sizes, and H_n is the property value of the nth residue in the sequence.

The Fourier power spectrum (FPS) of the sequence is calculated from a frequency of 2/(total window size) to 1/2 in a variable increment as follows:

$$I_{(j,v)} = \left[\sum_{n=j-w_l}^{j+w_r} (\bar{H}_w - H_n) \exp(2\pi i \cdot v \cdot n) \right]^2 \Big/ (w_l + w_r + 1)^2$$

where H_n is the value of the physical property of the nth residue, $\bar{H}_w$ is the average value over the window being sampled, w_l and w_r are the sizes of the left and right windows, respectively, and v is the frequency. End effects were handled similarly to the windowed averaging calculations with adjustment of the Fourier window at the ends to include only actual sequence values. The spectrum was normalized to the square of the total window size. The mean physicochemical value of the residues in the Fourier window were subtracted from individual values to remove the origin peak of the spectrum.

The α-helical region of the FPS was defined as the region with reciprocal repeat frequency between 1/3.0 and 1/4.5 (80–120°), and the β region of the transform was defined as the region between 1/2.4 and 1/2.0 (160–180°). These ranges encompass those frequently observed in protein structures (Blundell and Johnson, 1976).

The calculation of the amphipathic helical level (AHL) was developed as a measure of the interfacial level adopted by a helix along its axis based on the distribution of the physicochemical properties. The AHL analysis was calculated at each window position by a series of steps, and the final result was reported at the central window positions. In the first step, the FPS was calculated for the helical range of the transform. Next, the vector sum of the physical property values was calculated utilizing the frequency of the FPS maximum as the repeat frequency of the helix at each window position. A plane of sign inversion was next positioned parallel to the helical axis and perpendicular to the dipole vector at different levels within the helix, dividing the helix into two regions (see Fig. 8 *B*). The physical property values for residues within the current window were summed on each side of the interface plane, and the difference between the two sides was determined for each position of the plane as follows:

$$\mathrm{AHL}_{(n,l)} = \sum_{j=n-w_l}^{n+w_r} H_j \{ \sin[(v \times j) - \mu - 0.5\pi] - l \} / |\sin[(v \times j) - \mu - 0.5\pi]\} - l|,$$

where μ is the dipole angle, v is the repeat angle determined from the FPS, l is the sampled level from ±1 helical radius, w_r and w_l are the window sizes, and n is the residue number of the current window center. When all the positions of l (from +1 to −1 radius) had been sampled for each window position, the values at each position were normalized. Since most hydropathy scales assign a different sign to hydrophobic and hydrophilic values, the maxima reported by this analysis occur at the interface between hydrophilic and hydrophobic amino acids.

Established methods of structure prediction

The program SEQ by Teeter and Whitlow (1988) was the first method used for determining the secondary structure potentials for α-helix, β-sheet, β-turn, and random coil. This program is based on the method and uses the probability tables of Chou and Fasman (1978) or Levitt (1978) to assign secondary structures constrained to a fixed composition. This program also tests potential α-helical segments for amphipathic character by examining the residue hydropathies for a helical (3.6 residues per turn) period. The total helical probability is determined by the addition of an empirical constant scaled relative to the amphipathic character of the potential segment, thereby increasing the probability of helical segments demonstrating amphipathic character.

In the second method, secondary structure predictions for α-helix, β-sheet, and β-turn were carried out in unconstrained fashion using the method originally developed by Chou and Fasman as implemented in the program PREDICT (Prevelige and Fasman, 1989), utilizing the probability values derived from 64 proteins.

The third method of secondary structure predictions used the information theory based method implemented in the GOR algorithm (Garnier et al., 1978; Gilbrat et al., 1987). Extended, turn, and coil decision constants were adjusted relative to a fixed helical decision constant of 100 to vary the secondary structure composition of the final assignment.

The basic statistical method of Gascuel and Golmard (1988), as implemented with their GGBSM algorithm, was the fourth method used to assign helix, extended, and coil secondary structure states to the protein. The parameter N_h was fixed at 1.0, and the values of the N_e and N_c parameters were varied in a manner similar to the GOR method to control the composition of the assignment.

General methods

The amino acid sequence of apoA-I used in the primary sequence analysis methods was that originally determined by Brewer et al. (1978), with the Gln to Glu corrections made as suggested by the gene sequencing (Karathanasis et al., 1983). The sequences of apoE-3 and apoA-IV were those determined by Rall et al. (1982) and Karathanasis et al. (1986), respectively. The sequence of apoLp-III used in this study was supplied by M. A. Wells (personal communication, 1992).

The GGBSM program (implemented in Pascal) and the PREDICT program (implemented in the C programming language) were both executed on an IBM PC/AT compatible computer. All other programs were written in VAX/VMS Fortran and executed on a MicroVax II computer (Digital Equipment Corporation, Maynard, MA).

Image data files produced by the homology and the CAPPS programs were displayed by the Spider Image Analysis System (Frank et al., 1981). Statistical analysis of the homology, physicochemical, and secondary structure methods were carried out with the RS/1 program (BBN Software Products Corp., Cambridge, MA).

RESULTS

Characterization of apolipoprotein-lipid complexes

Complexes of apoA-I/DMPC and apoE-3/DMPC were shown by negative staining electron microscopy to be

TABLE 1 **LINEQ analysis of apoprotein CD spectra**

Data set	Helical content	Sheet content	Turn content	Data range	Scale factor	NRMSD
A. ApoA1/DMPC	0.75	0.10	0.15	215–240	0.82	0.016
B. ApoA1/DMPC/CO*	0.76	0.05	0.19	205–240	1.10	0.037
C. ApoA1/DPPC*	0.59	0.15	0.26	204–240	1.16	0.029
D. ApoE/DMPC	0.56	0.21	0.22	213–240	1.03	0.021
E. ApoE/DMPC‡	0.66	0.00	0.34	207–240	1.06	0.068

A comparison of the LINEQ results of the experimental data (A and C) and previously published spectra (B, D, and E) of similar lipid/protein systems.

* Spectra from Lux et al. (1972).

‡ Spectra from Roth et al. (1977).

100 Å diam disk-shaped particles. These particles are similar in morphology to nascent HDL disks (Atkinson and Small, 1986). Deconvolution of the observed CD spectra was carried out using LINEQ (Mao and Wallace, 1984) with all of the available data sets; however, the GREFAS reference data set (Greenfield and Fasman, 1969) provided the best fit to the observed data. The final secondary structure composition analysis for each protein in a lipid environment was taken as the mean of the results obtained from this CD study and previously published data (shown in Table 1). In the case of apoA-I, the composition was determined to be 70% α-helix, 10% β-sheet, and 20% turn and random structure, and for apoE-3 61% α-helix, 11% β-sheet, and 28% turn and random structure.

Sequence homology

Homology calculations were carried out within and between apoA-I and apoE-3. In addition, calculations were carried out with the apoA-IV sequence to refine the previously identified 11 amino acid internal repeat within this group of related proteins. A window size of ± 5 residues was used as it is the optimum size for the characterization of an 11 amino acid repeat.

The homology matrices for these comparisons (not shown) showed multiple significant diagonal maxima at intervals of 11 residues and thus confirm the fundamental repeat to be 11 residues. These 11-mer repeats are distributed throughout most of the three sequences with the exception of the NH_2-terminal regions (1–50 in apoA-I, 1–40 in apoE-3, 1–60 in apoA-IV). Local clusters of similar residues were also evident in the analysis of the repeats in apoE-3 (185–205, 279–295) and in apoA-IV (350–370).

Determination of a structural consensus sequence

The 11-mer repeats identified in apoA-I, apoE-3, and apoA-IV were considered initially as one group, and a scatter plot of the distribution of residues at each position (Table 2) indicated that positions 1, 3, 7, 8, and 9 contain at least two significant clusters of amino acids. Positions 2 and 11 showed little clustering and the remaining positions 4, 5, 6, and 10 showed only one dominant amino acid per position.

Correlation analysis performed on the residue bulk, pKi, hydropathy, and the secondary structure probabilities of the amino acids specified by the 34 11-mer repeats from apoA-I and apoE-3 indicated weak correlations between the properties of the residues at positions 1, 8, and 9. Since the two residue clusters at position 9 (Arg and Lys) shared similar values in the scales examined no significance could be drawn for the consensus sequence from a correlation at that position. However, the clusters at positions 1 and 8 were between residues with dissimilar properties (Pro and Ala at position 1; Ala and Glu at position 8). This suggests a minimum of two types of 11 amino acid repeat delimited by the residues at positions 1 and 8. The previous assignments of Li et al. (1988) to two similar types of 11-mer repeats that included a different consensus residue at position 1 and 8 correlates with these data.

An initial consensus sequence having the dominant amino acid at each position was derived from the A/B repeat assignments of Li et al. (1988). The repeats within apoA-I, apoE-3, and apoA-IV were then reassigned as either type A or type B repeat using the average homology calculated for the repeat against this initial consensus sequence.

The use of mutation based criteria in the homology calculations resulted in several of the assignments changing from the type A or B classifications of Li et al. (1988). In addition, Monte Carlo background correction identified several regions that did not show significant homology above the background level. Also, there were several regions that showed equal identity to the A and the B repeats. These regions were not used in the final determination of the consensus sequences. Refined consensus sequences were formulated from these new assignments in a similar manner and the homology of each region to the new consensus sequences was determined. Iterative consensus sequence refinement was repeated until only minor changes resulted between subsequent iterations.

TABLE 2 **Frequency of amino acid occurrence at each position of an 11 AA repeat in apoA-I, apoE-3, and apoA-IV**

	Position											
AA	1	2	3	4	5	6	7	8	9	10	11	Hydropathy*
Ile						1				6		1.38
Phe	1	1	1			6		1		1		1.19
Val		3	5			6		2	1	5	3	1.08
Leu	1	8	13		1	27				28	2	1.06
Trp		1					1					0.81
Met			5			3	1			2		0.64
Ala	10	1	15	3	5	1	2	9	2	5	6	0.62
Gly	2	1	3	3	3		2	5	1	1	8	0.48
Cys							1					0.29
Tyr		7				3						0.26
Pro	17											0.12
Thr	1	3	3	1	1	3	2	2		1	2	−0.05
Ser	3	1	2	4	3		2	1	4		5	−0.18
His	1	4		1			2		3			−0.40
Glu	3	3	1	24	23		3	11	3		4	−0.74
Asn		4		1	1		1		3		4	−0.78
Gln	4	4	1	4	2	1	10	8	7		1	−0.85
Asp		2		8	6			3			3	−0.90
Lys	4	2	1	1	4		7	6	12		5	−1.50
Arg	1	3	1	2	2		17	3	16		6	−2.53

* Hydropathy values taken from "Optimized Hydropathy Scale" of Eisenberg (1984).

The next phase of the refinement used a binary substitution matrix based on bulk and hydropathy (Bordo and Argos, 1990, 1991) that may more accurately represent structurally equivalent residues. Two additional iterations of refinement were necessary to achieve a stable assignment that did not exchange A and B regions between rounds. The final consensus sequence is shown at the top of Table 3.

An image representation of the conservative substitution based and background corrected homology within apoA-I, apoE-3, and apoA-IV to a tandem repeat of the refined A and B consensus sequences is shown in Fig. 1. The final assignment of the repeats in apoA-I, apoE-3, and apoA-IV to either the A or the B type consensus sequences is shown in Table 3. Regions with equal homology to the A and B repeats are reported in the positions originally assigned by Li et al. (1988).

Physical chemical properties

The local averages of the physicochemical properties were calculated for apoA-I and apoE-3. Hydrophobic moment analysis was performed with a window of ±5 residues, at a frequency of 1/3.6 for α-helix and a frequency of 1/2.0 for β-sheet. Finally, one-dimensional power spectra were calculated for the physical properties; however, significant results were observed only for hydropathy based scales.

ApoA-I

Fig. 2 *A* illustrates in image format many of the physical properties of apoA-I, including secondary structure probabilities. The hydropathy scales used in this study included the "PRIFT" scale (Cornette et al., 1987), which is optimized for the detection of helices, and the optimized hydropathy scale (OPTHYD) of Eisenberg (1984), which is a combination of other scales. The use of both of these scales identified regions with high overall hydrophobicity (13–22, 41–43, 214–235) and regions with hydrophilic character (1–9, 103–108). Use of the PRIFT scale identified additional hydrophobic regions (158–161, 173–178) and the OPTHYD scale suggested additional hydrophilic regions (77–84, 128–135, 147–152).

Average residue bulk calculation suggests several regions where clusters of smaller resides occur (27–40, 82–84, 147–153, 180–191). These regions may correlate with areas of conformational flexibility such as turn or hinge regions. Regions with a large mean residue bulk are also apparent (12–20, 42–49, 95–104, 116–119, 216–226).

Fourier analysis of the hydropathy of apoA-I (Fig. 2 *A*) shows prominent maxima at an α-helical frequency (80–120°) for most of the sequence. However, several regions (1–14, 30–45, 107–115, 150–158, 180–210, 225–230) lack significant maxima in this range. Significant maxima in the β repeat region of the transform are evident (35–45, 90–95, 105–115, 135–145). The Fourier calculations presented were carried out using the PRIFT hydropathy scale; however, the OPTHYD scale showed only minor differences from these results.

Variation of window size had a dramatic effect on the Fourier analysis. The larger window sizes result in a

TABLE 3 Assignment of consensus sequence to apoA-I, apoA-IV, and apoE-3

		Derived consensus sequence		
1	A	11	1	B 11
P L A E E L R A R L R			A Q L E E L R E R L G	
		Human apolipoprotein A-I		
			44 L K L L D N W D S V T	54 (08)*
			55 S T F S K L R E Q L G	65 (47)
66 P V T Q E F W D N L E	76 (41)			
77 K E T E G L R Q E M S	87 (25)*			
			88 K D L E E V K A K V Q	98 (33)‡
99 P Y L D D F Q K K W Q	109 (13)‡		110 E E M E L Y R Q K V E	120 (21)
121 P L R A E L Q E G A R	131 (26)		132 Q K L H E L Q E K L S	142 (34)
143 P L G E E M R D R A R	153 (61)		154 A H V D A L R T H L A	164 (41)
165 P Y S D E L R Q R L A	175 (48)		176 A R L E A L K E N G G	186 (39)
			187 A R L A E Y H A K A T	197 (14)*
			198 E H L S T L S E K A K	208 (13)
209 P A L E D L R Q G L L	219 (27)			
220 P V L E S F K V S F L	230 (25)		231 S A L E E Y T K K L N	241 (31)
		Human apolipoprotein A-IV		
40 A L F Q D K L G E V N	50 (12)‡			
51 T Y A G D L Q K K L V	61 (25)*			
62 P F A T E L H E R L A	72 (44)		73 K D S E K L K E E I G	83 (33)
			84 K E L E E L R A R L L	94 (37)‡
95 P H A N E V S Q K I G	105 (33)		106 D N L R E L Q Q R L E	116 (32)
117 P Y A D Q L R T Q V N	127 (44)		128 T Q A E Q L R R Q L D	138 (40)‡
139 P L A Q R M E R V L R	149 (36)			
150 E N A D S L Q A S L R	160 (30)			
161 P H A D E L K A K I D	171 (44)		172 Q N V E E L K G R L T	182 (40)‡
183 P Y A D E F K V K I D	193 (35)		194 Q T V E E L R R S L A	204 (46)
205 P Y A Q D T Q E K L N	215 (19)		216 H Q L E G L T F Q M K	226 (30)
227 K N A E E L K A R I S	237 (47)		238 A S A E E L R Q R L A	248 (55)
249 P L A E D V R G N L K	259 (59)		260 G N T E G L Q K S L	269 (32)
270 A E L G G H L D	277 (30)		278 Q Q V E E F R R R V E	288 (56)
289 P Y G E N F N K A L V	299 (33)		300 Q Q M E Q L R Q K L G	310 (52)
311 P H A G D V E G H L S	321 (32)			
322 F L E K D L R D K V N	332 (24)*			
		Human apolipoprotein E-3		
			62 A L M D E T M K E L K	72 (20)*
			73 A Y K S E L E E Q L T	83 (27)
84 P V A E E T R A R L S	94 (61)		95 K E L Q A A Q A R L G	105 (21)
			106 A D M E D V C G R L V	116 (33)
117 Q Y R G E V Q A M L G	127 (19)‡		128 Q S T E E L R V R L A	138 (45)‡
			139 S H L R K L R K R L L	149 (33)*
			150 R D A D D L Q K R L A	160 (26)‡
			161 V Y Q A G A R E G A E	171 (02)*
			172 R G L S A I R E R L G	182 (42)
183 P L V E Q G R V R A	192 (27)‡		193 A T V G S L A G Q	201 (10)
202 P L Q E R A Q	208 (29)			
209 A W G E R L R A R	217 (50)*			
218 M E E M G S R T R	226 (35)		227 D R L D E V K E Q V A	237 (40)
			241 A K L E E Q A Q Q I R	251 (31)*
252 L Q A E A F Q A R L K	262 (37)*			
267 P L V E D M Q R Q W A	277 (23)‡		278 G L V E K V Q A A V G	288 (38)

Numbers in parentheses denote the average Monte Carlo corrected homology of the region to the A or B consensus sequence. *A/B assignment differs from Li et al. (1988). ‡Homology to A and B are the same. Shown in position of Li et al. (1988).

smoothing of the plot and a focusing of the maxima. At a window size of ±11 residues, the breaks in the helical maxima around residues 10, 40, 190, and 230 remained.

The balance of the sequence displayed maxima near 100°. When the window size was reduced to ±4 residues (or ~1 helical turn in each direction), few significant extended maxima appear. A window size of ±5 residues

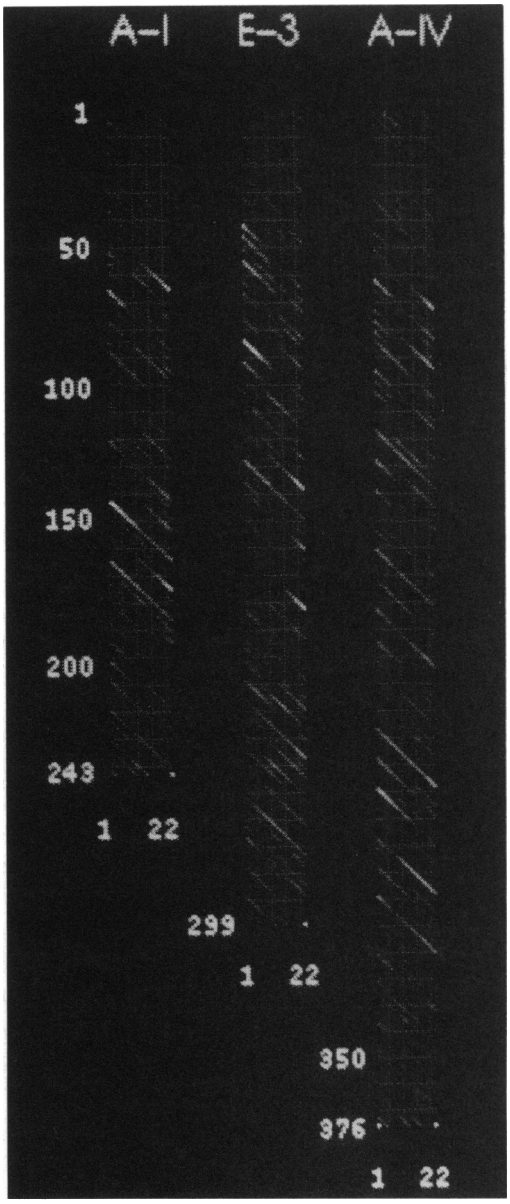

FIGURE 1 Consensus repeats in apoA-I, apoE-3, and apoA-IV. An image representation of the homology comparison matrix of an A/B tandem repeat of the A-I/A-IV/E-3 consensus sequence (PLAEELRARLR-AQLEELRERLG) against the sequences of apoA-I, apoE-3, and apoA-IV. Homologies were determined using a similarity matrix derived from data by Bordo and Argos (1990), and Monte Carlo background correction was applied. The consensus sequence maps to the horizontal dimension of each matrix, and the protein sequence corresponds to the vertical dimension. Each individual pixel (i, j) represents the average homology for a window of 11 amino acids centered at residue i in the first protein and j in the other. Images are displayed using the program SPIDER and are shaded with the BW8 scale.

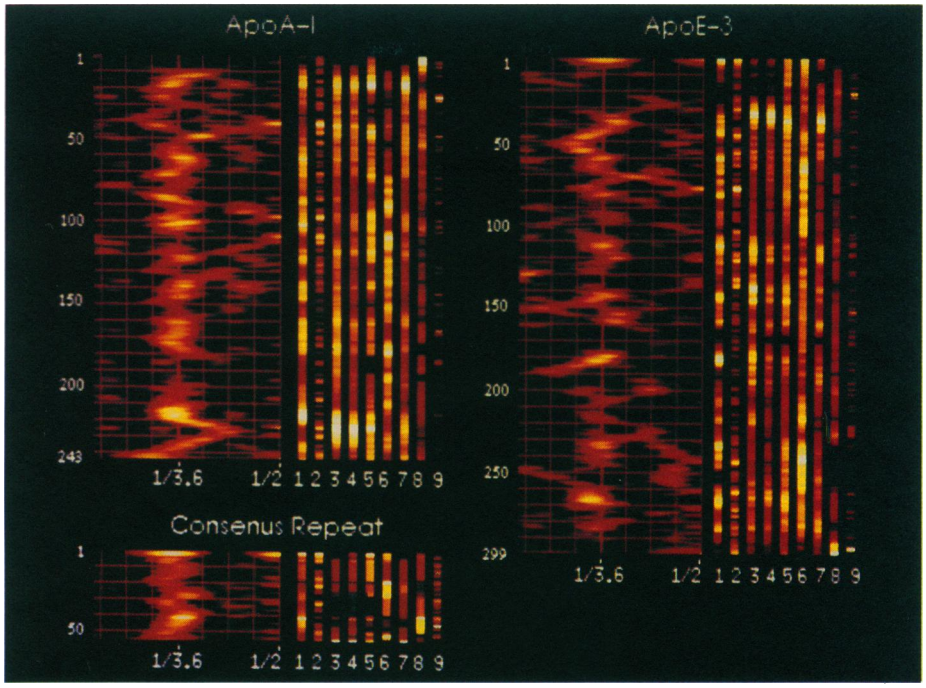

FIGURE 2 Image summaries of the physicochemical profiles of human apoA-I, apoE-3, and the refined consensus sequences. Image summaries of apoA-I, apoE-3, and an AABBA repeat of the refined A-I/E-3/A-IV consensus sequence are shown. All calculations except the β-turn frequency are calculated with a total window size of 11 and displayed at the position of the central residue of the window. The properties shown for each sequence from left to right are: the Fourier power spectrum of PRIFT hydropathy from 1/5.5 to 1/2, hydrophobic moment at 100° (1), hydrophobic moment at 180° (2), mean PRIFT hydropathy (3), mean OPTHYD hydropathy (4), mean residue bulk (5), mean α-helix probability (6), mean β-sheet probability (7), mean β-turn probability (8), and β-turn frequency (9). The vertical axis represents position along the sequence (amino acid number) in all images. Grid lines are drawn across the horizontal axis of the FPS at intervals of 1/20.0 with a bold grid line at a frequency of 1/3.6. Images are displayed with the program SPIDER using the METHEA scale that corresponds to the temperature of a glowing body.

represents the best compromise for detection of local variations within the structure. This size window corresponds also to the average helical length in proteins and is the size of the homology repeat unit.

ApoE

The image representations of the sequence properties of apoE-3 are shown in Fig. 2 *B*. Calculations were carried out in an identical manner to that described for apoA-1.

The average hydropathy calculation identified clusters of hydrophobic residues (28–43, 111–118, 136–146, 176–182, 193–195) and hydrophilic residues (1–24, 167–171, 203–210, 222–231, 242–251, 292–299) with both the OPTHYD and the PRIFT scales. The PRIFT scale (Cornette et al., 1987) also identified the region 75–80 as hydrophilic in the apoE-3 sequence. Average bulk analysis identified several regions of low bulk (99–108, 164–176, 221–223) and regions of high bulk (32–44, 262–272).

Fourier analysis of apoE-3 again exhibited maxima in the helical range as the major feature. However, the NH_2-terminal region from residues 10–35 is devoid of maxima with helical frequency. Other regions (160–175, 190–200, 245–255, 285–299) did not exhibit maxima in the helical range of the FPS. Maxima were also seen in the β range of the FPS (75–80, 130–140, 235–265).

Consensus sequences

The image analysis of an A-A-B-B-A consensus repeat sequence is shown in Fig. 2 *C*. This combination allows examination of the possible junctions of the two 11-mers (A-A, A-B, B-B, B-A) as well as a complete A-B repeat unit.

Both the A and B repeat sequences exhibit high helical probabilities. The repeat sequences also exhibit maxima in the helical range of the FPS, with the B repeat having a higher maximum. The helical probability is low at all four repeat junctions. However, helical probability is lowest at a B-A junction. This junction also contains the highest turn probability due in part to the lead proline which can function as a helix breaker. Sheet probability remains low except at the junction between an A and a B repeat; however, it remains lower than the helical probability at this point.

Strong maxima in the helical range of the FPS, secondary structure potentials, and secondary structure prediction methods suggests that the A and the B consensus

sequences are likely to form helices. When multiple copies of the A or the B sequences are juxtaposed, turns are predicted only at the junction of a B sequence followed by an A sequence. An A sequence followed by another A sequence weakly predicts turn and an A sequence followed by a B sequence does not predict turn at the A-B junction.

Modeling of a single 22 amino acid segment composed of an A repeat followed by a B repeat shows that this segment if folded into a helix would form a very amphipathic structure. The A and B repeat remain in register with each other when viewed down the helical axis as shown in the molecular graphics model in Fig. 3 *A*. A single hydrophobic face followed by an interfacial region containing all of the Arg residues would be generated. Opposite the hydrophobic face is a hydrophilic region containing all of the Glu residues and two Ala residues. The location of these residues is shown in a helical wheel representation (Fig. 3 *C*). A Gln residue located at the edge of the hydrophobic face is apparent also in this projection. It has been suggested that a Glu residue in this position may be important in the activation of LCAT (Anantharamaiah et al., 1990). When the helix is viewed from the side, these repeats show a longitudinal segregation of alternating positively and negatively charged regions. Fig. 3 *B* shows two antiparallel A/B segments in which this charge segregation is apparent.

Secondary structure prediction

The secondary structure predictions were carried out using the four established methods for apoA-I and apoE-3. Scaling parameters for SEQ, GGBSM, and GOR were varied such that the predictions were constrained in a range of ±5% of the compositions determined by CD. Multiple executions of each program allowed sampling of the predicted conformations. For apoA-I, the ranges were 65–75% α-helix, 5–15% β-structure, and 15–25% turn/random. For apoE-3, the ranges 56–66% α-helix, 6–16% β-structure, and 23–33% turn/random were used. The sum of the conformational states in each prediction was constrained to 1.0 in each of the algorithms.

The multiple predictions from each individual method were reduced to a table of the frequency at which each state (helix, β, turn, or random) was predicted for each sequence position. A summary of the predictions from these methods is shown in Fig. 4 *A* for apoE-3 and Fig. 4 *B* for apoA-I. These figures show the consensus prediction of each of these methods together with a summary of the sequence homology to the A or the B consensus sequences.

Integration of prediction methods

An integrated assignment of secondary structure was carried out in a multi-step process. The initial criterion for assignment was the concurrent prediction of the same secondary structure state by all four methods at a frequency of 90% or greater on variation of the constraints.

For the next series of steps, the criteria were relaxed requiring only three of the four methods to predict a state at a 75% frequency. However, before an assignment was made, physicochemical or probability data, subjected to significance testing, was required to support the assignment. Specifically, the assignment was initiated by assigning turn residues at positions where three of the four methods predicted turn, and the β-turn frequency was $\geq$2.0 standard deviations above the average level of turn frequency. Amphipathic α-helix was assigned where three of four methods predicted α-helix, and a maximum occurred in the helical region of the FPS at $\geq$2.0 standard deviations above the average for that region of the FPS. Amphipathic β-structure was assigned when three of four methods predicted β-structure, and a maximum occurred $\geq$2.0 standard deviations above the average in the β region of the FPS. In addition, turn, α-helix, and β-structure were assigned when three of the four methods predicted that state, and there was a maximum $\geq$ 2.0 standard deviations above the mean of their respective probabilities.

This series of steps was repeated with the criteria progressively relaxed during subsequent iterations. The criteria used in this phase of the assignment are summarized in Table 4, steps 1–25.

In the next phase of the assignment (Table 4, step 26), homology data between the sequence of apoA-I (or apoE-3) and the sequence of apoLp-III were used to assign regions of secondary structure. The central residue of a ±3 amino acid window in apoA-I and apoE-3 was assigned the secondary structure observed in the crystal structure of apoLp-III (Breiter et al., 1991) when the following conditions were met: (*a*) the homology at that residue position was $\geq$3 SD above the Monte Carlo calculated background; (*b*) when multiple sites of significant homology to apoLp-III occurred, each site predicted the same secondary structure state; and (*c*) the secondary structure state of the matching position in apoLp-III was defined unambiguously in the crystal structure.

Homology matching was also carried out in a similar manner against the sequence of the NH_2-terminal domain of ApoE-3 (step 27), for which the crystal structure has been determined (Wilson et al., 1991). Homology data for this apoE-3 fragment was not included in the corresponding region of the apoE-3 prediction.

This step was followed by a similar assignment (Table 4, step 28) to residues that matched the theoretical structure of the consensus sequence. Since the structure of the terminal residues of the type A and B repeats was predicted to be dependent on the repeat type and the type of the adjacent repeat, the homology was restricted to assignments that excluded the three terminal residues of each repeat.

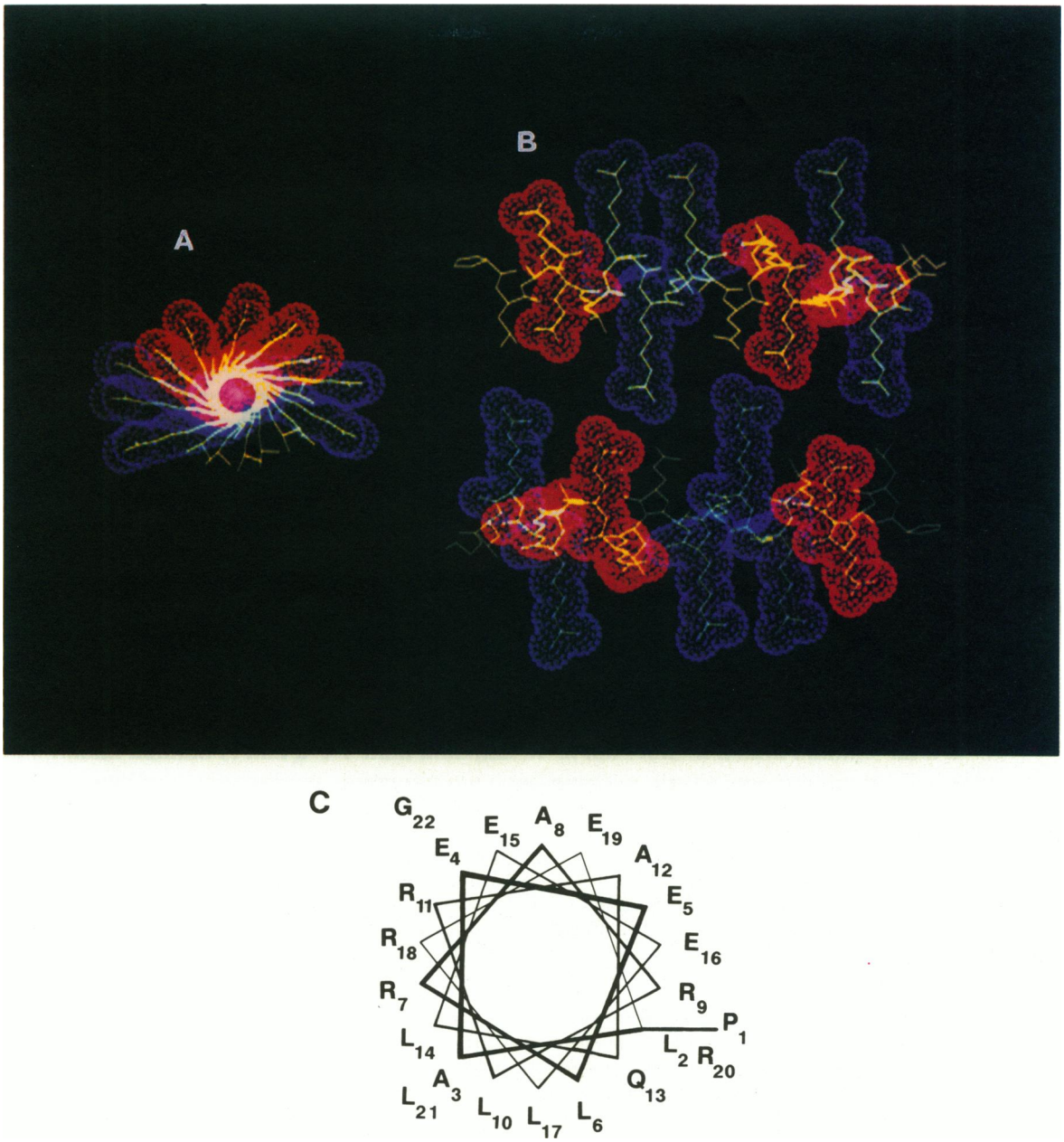

FIGURE 3 Molecular model of the an A/B tandem repeat of the A-I/A-IV/E-3 consensus sequence. The model for the consensus sequence was created with the program FRODO (Jones, 1982). Acidic residues are shown in red and basic residues in blue. Neutral residues are shown in green. Helical structures were created using the phi-psi angles of −57.0° and −47.0°, respectively. Regularization of the structure was carried out with the Herman's regularization routine available in FRODO. *A* shows a view down the axis of the consensus sequence when folded into a helix. *B* shows a helical bundle composed of two antiparallel A-B consensus sequences when viewed from the side, illustrating the noncovalent interactions possible between residues on adjacent helices. *C* shows a helical wheel of the consensus sequence.

The final steps of the assignment were carried out using a procedure similar to the first phase but without the inclusion of the secondary structure prediction methods. Physicochemical property levels were reduced in two steps to minimum significance. These levels are shown in Table 4, steps 29–36.

The assignment history from the integrated assignment protocols are shown in Fig. 5 for apoE-3 and Fig. 6 for apoA-I. Regions referred to as "Turn" include all B-Turn types as well as other defined loop structures. No attempts to distinguish between these classes have been made. These regions have been enumerated only at the central positions since the terminal positions share the properties of the structures abutting the turn. The final phase of integrated assignment of apoA-I resulted in 136 residues with unique assignments, 91 residues assigned to multiple states, with 16 residues that were completely undefined. For apoE-3, this integrated approach resulted in 149 uniquely assigned residues, 120 residues with multiple assignments, and 30 residues that were undefined.

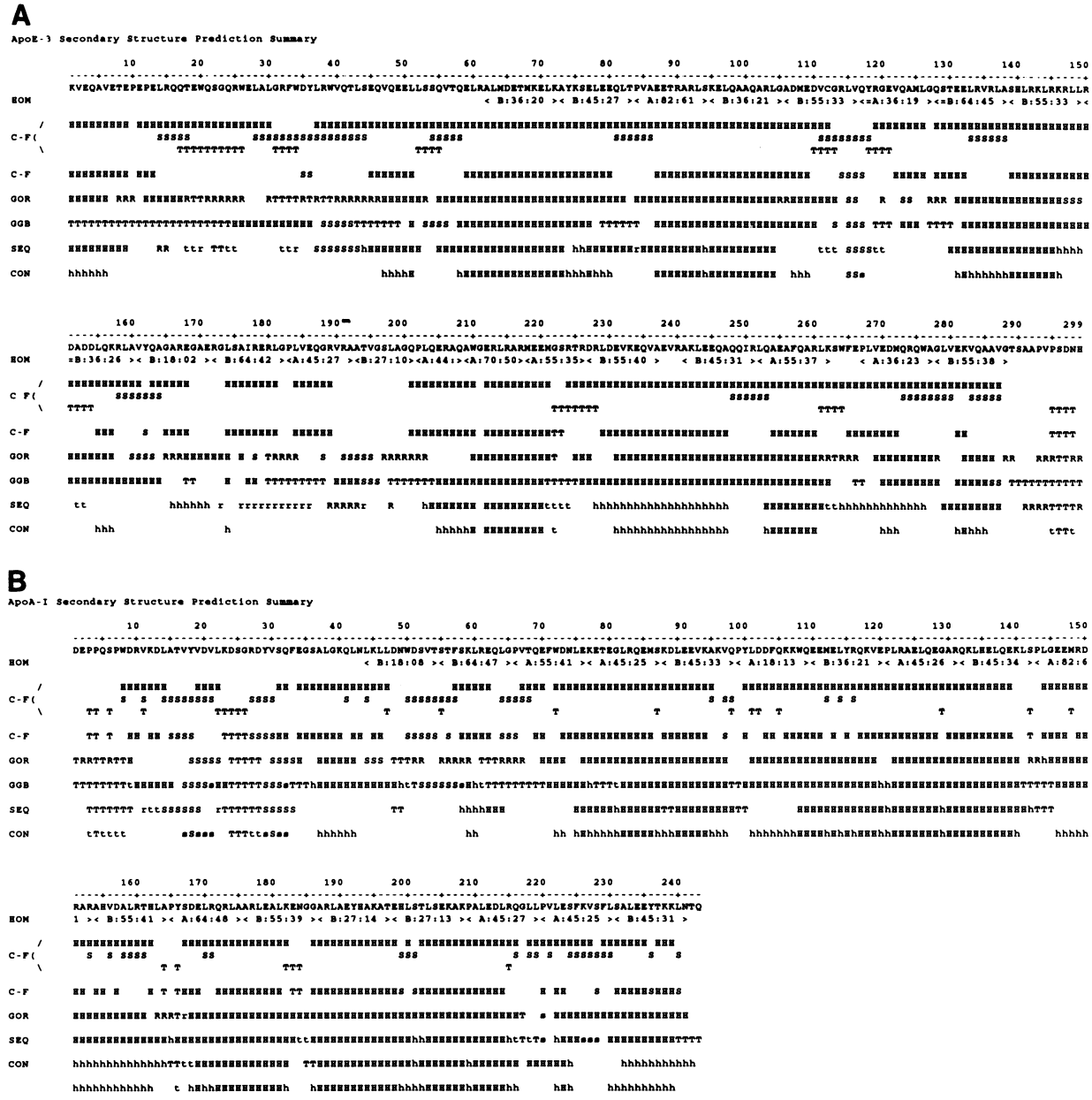

FIGURE 4 Secondary structure prediction summary for apoA-I and apoE-3. The results of the Chou-Fasman analysis (C-F) shown were calculated using the 64 protein database. Completely undefined regions were considered to be predicted as random coil in the frequency table derived from this analysis. For the Garnier et al. (1978) analysis (GOR), the results shown are a consensus prediction from seven runs for both apoE-3 and apoA-I. The results of the Whitlow and Teeter method (SEQ) shown were the consensus prediction from 30 runs for apoE-3 and 20 runs for apoA-I. The precedence of structure assignment was varied also for both proteins between turn/helix/sheet/coil and turn/sheet/helix/coil. Empirical helical probability increases due to amphipathic characteristics was both allowed and disallowed for both proteins. The consensus of four runs of the Gascuel and Golmard (GGB) algorithm are shown for each protein. The final consensus predictions (CONS) result from a comparison of the four methods. For all results shown when ≥90% of the predictions within a method predicted the same state a capital letter is shown. Lower case letters are shown when the prediction was ≥70%. The homology to the derived consensus sequence is shown also on these plots (HOM). The data is represented as ⟨repeat type:uncorrected similarity:background corrected similarity⟩. Similarity was determined using the Bordo and Argos (1990) derived substitution matrix.

DISCUSSION

Methods for increasing the accuracy of predictive methods by combination of multiple secondary structure prediction algorithms (Schultz et al., 1974; Matthews, 1975) and the integration of predictive methods with hydrophobicity data (Biou et al., 1988) or homology data (Nishikawa and Ooi, 1986) have been demonstrated. In this study, four secondary structure prediction algorithms, physicochemical data, and a generalized

TABLE 4 **Integrated prediction assignment protocol**

Step no.	Secondary methods		Physical methods	
	Match	Methods	Protocol	Min level
	%			
1	90	4	—	—
2	75	3	1	Mean + 2SD
3	75	3	2	Mean + 2SD
4	75	3	3	Mean + 2SD
5	75	3	4	Mean + 2SD
6	75	3	5	Mean + 2SD
7	75	3	6	Mean + 2SD
8	75	3	1	Mean + 1SD
9	75	3	2	Mean + 1SD
10	75	3	3	Mean + 1SD
11	75	3	4	Mean + 1SD
12	75	3	5	Mean + 1SD
13	75	3	6	Mean + 1SD
14	50	3	1	Mean + 1SD
15	50	3	2	Mean + 1SD
16	50	3	3	Mean + 1SD
17	50	3	4	Mean + 1SD
18	50	3	5	Mean + 1SD
19	50	3	6	Mean + 1SD
20	50	2	1	Mean + 1SD
21	50	2	2	Mean + 1SD
22	50	2	3	Mean + 1SD
23	50	2	4	Mean + 1SD
24	50	2	5	Mean + 1SD
25	50	2	6	Mean + 1SD
26	apoLp-III homology		7	Bkg Mean + 3SD
27	apoE-3 fragment homology		7	Bkg Mean + 3SD
28	A/B cons. homology		7	Bkg Mean + 3SD
29	—	—	2	Mean + 1SD
30	—	—	3	Mean + 1SD
31	—	—	1	7.5×10^{-5}
32	—	—	5	Mean
33	—	—	6	1.0
34	—	—	4	Mean + 1SD
35	—	—	2	Mean
36	—	—	3	Mean

Protocol 1: Turn is assigned from the β-turn frequency data; 2: helix is assigned from the helical range of the hydrophobic power spectrum (80–120°); 3: sheet is assigned from the β range of the hydrophobic power spectrum (160–180°); 4: turn is assigned from the β-turn probability; 5: helix is assigned from the helical probability; 6: sheet is assigned from the sheet probability; 7: homology: minimum level calculated relative to background for entire protein.

substitution-based homology scheme have been simultaneously applied to the soluble apolipoproteins apoE-3 and apoA-I.

The recent crystal structure of the NH_2-terminal domain (residues 1–191) of apoE-3 (Wilson et al., 1991) provided an opportunity to evaluate and refine the integration of these data in the structural assignment. In the assignment of apoE-3, the region of the sequence corresponding to this fragment was assigned without the aid of the known structure. This provided both for feedback used in the evaluation of our methodology and the potential detection of regions of the sequence whose structure may be dynamic and depend on the environment of protein (i.e., in aqueous solution or bound to a lipoprotein particle). The assignment of the remainder of the apoE-3 sequence and the apoA-I sequence used the detailed structure of the apoE-3 fragment directly in the homology protocol. A final structure assignment for apoE-3 and apoA-I has been carried out by examining the assignment history in the integrated procedures. Clues to the appropriate weighting of these data were obtained by examining the parameters that best correlate to the structure identified in the apoE-3 fragment.

ApoE-3 model

The integrated protocols resulted in a secondary structure prediction of 90% of apoE-3. A model (shown in Fig. 7 *A*) was derived after the assignment of the remainder of the sequence by detailed inspection of the properties of each region and consideration of the minimum segment lengths of structural elements. In addition, greater emphasis was given to the assignments made by the most stringent selection criteria. The secondary structure constraints derived from analysis of CD spectra of apoE-3/DMPC complexes provided additional guidelines.

The NH_2-terminal domain is predicted to begin with a small amphipathic helix (residues 1–6). This is followed by a region of random coil (7–24) that was assigned because of dominant turn predictions throughout the region and the lack of a significant maximum in the FPS. Residues 25–34 have been assigned β-structure to complete a partially predicted segment in this area. The next region (35–42), which was predicted strongly by the statistical methods as β-structure, exhibited a strong helical amphipathicity and a significant homology to a helical region of apoLp-III. For these reasons, helical structure was assigned to residues 35–40.

The region between residues 43 and 159 was predicted to be predominantly helices of varying amphipathicity punctuated by turns. However, the turns in this region were poorly defined. One turn was assigned in a region (53–54) that was ambiguously defined by these protocols. A second turn has been assigned at a B/A repeat junction (84–85) that was predicted as helix by the integrated methods but judged to be the best candidate to subdivide what would otherwise be a 50 residue helix. An additional turn was assigned tentatively at a position (105–106) undefined by the assignment protocol where disruptions in the helical region of the FPS occur. A region of random coil has been assigned between two strongly predicted turns (118–128) that were separated by a region containing moderate helical amphipathicity and a strong β-structure prediction by statistical methods.

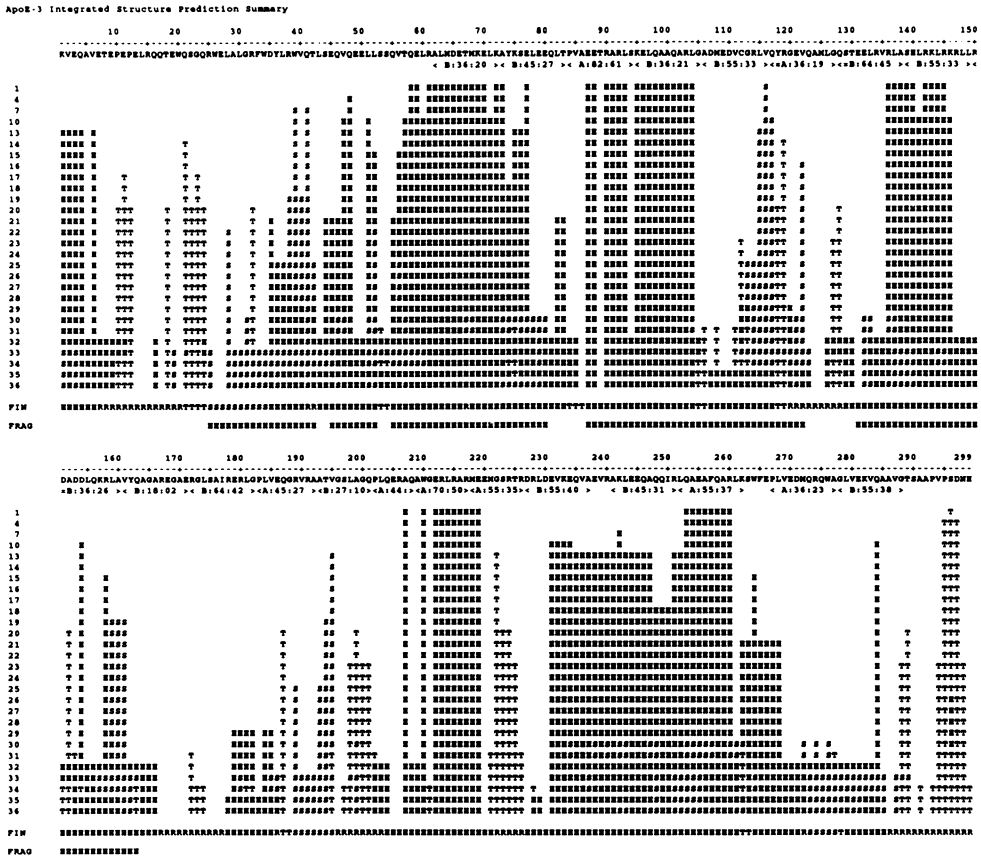

FIGURE 5 Apolipoprotein E-3 secondary structure assignment history. Assignment levels (1–36) correspond to those shown in Table 4. The final assignments (FIN) were made as described in the discussion. Helical regions observed in the crystal structure of the N-terminal apoE-3 fragment (Wilson et al., 1991) are shown (FRAG).

The well-defined amphipathic helix (55–83) predicted in this region would require a dislocation or twist around residue 67 to optimally align the hydrophobic face of the helix. The helix that follows (87–104) was strongly predicted by secondary structure potentials but did not exhibit a strong amphipathic character in the FPS. The region between 107 and 117 was predicted as helix despite strong predictions by the secondary structure prediction algorithms of β-sheet. A maximum in the helical range of the FPS and the model building supported the existence of helix in this region. These ill-defined segments may represent regions within apoE-3 that may have a dynamic structure determined by the environment. The region (129–166) is predicted as amphipathic helix and contains the receptor binding domain of apoE-3 (Weisgraber et al., 1983). This region contains a cluster of basic residues located in the center of 3 B-type helical repeats in this model.

The next region (167–177) that was ambiguously defined exhibited no significant maxima in the FPS and contained a cluster of residues with low bulk, and high hydrophilicity has been assigned random coil. These properties suggest that this area may function as a hinge region. This segment is followed by an amphipathic helix (178–185), several residues of turn that extend to residue 188, and a β-structure between residues 189–195. Following this strand and extending to residue 203, a region of random coil is predicted. A region of weakly amphipathic helix is strongly predicted in the next region of the molecule (204 and 221). A random coil/turn region is predicted between residues 222 and 226.

The COOH-terminal domain of the molecule is composed of a 35-residue region of moderately amphipathic helix (227–261) that was uniquely predicted from a very early stage in the integrated protocol. This is followed by a weakly predicted turn (262–263) and a helix (264–271) that exhibits a very strong maximum in the helical region of the FPS. A short region of β-structure is predicted next at residues 273–277. A second short helix is predicted (279–285) that is moderately amphipathic by the same criterion. The carboxyl terminus (286–299) exhibits no amphipathic character in the FPS and appears to be unstructured from the strong prediction of turn or coil residues.

The final assignment of the sequence of apoE-3 resulted in 193 residues (65%) helix, 22 residues (7%)

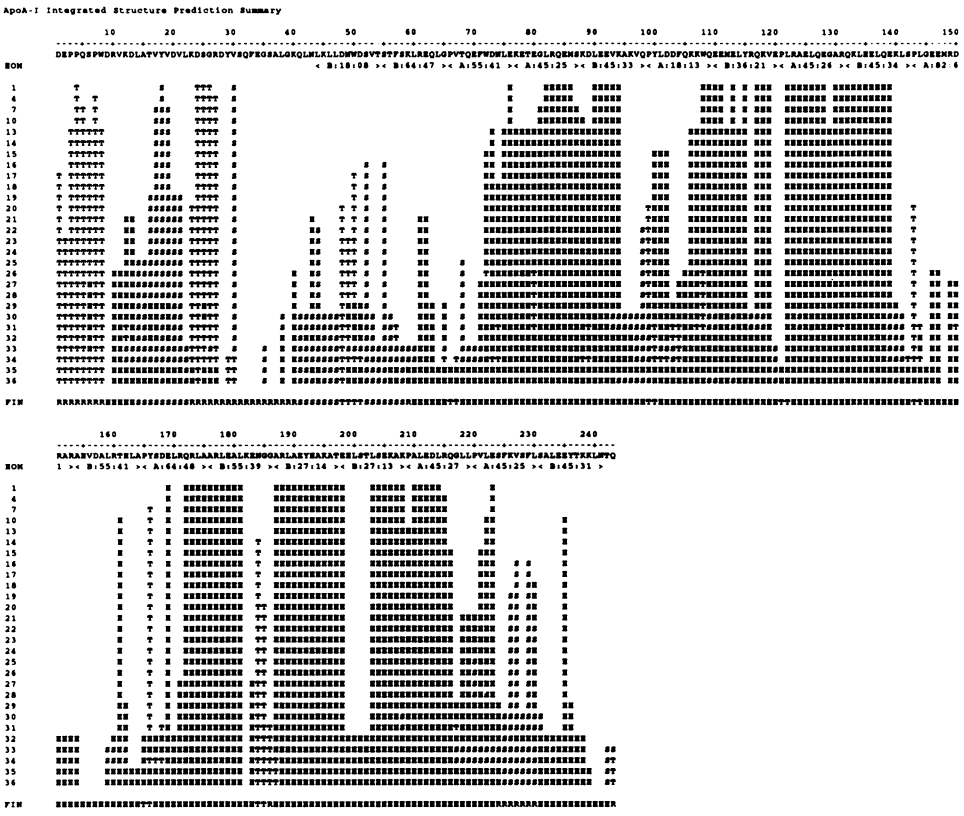

FIGURE 6 Apolipoprotein A-I secondary structure assignment history. Assignment levels (1–36) correspond to those shown in Table 4. The final assignments (FIN) were made as described in the Discussion.

sheet, and 84 residues (28%) to random/turn. The CD-derived constraints for this protein were 61% helix, 11% sheet, and 28% turn/random.

Enzymatic cleavage in the region (194–234) has shown that apoE-3 is folded in two independent domains (Wetterau et al., 1988). In this study, residues from 167–226 were identified as a region of low amphipathic character, containing segments of low bulk and segments with weak secondary structure potentials. These features are consistent with the structure of a hinge region between individual domains.

In this model, 118 residues (85%) are predicted to be in the same conformation as the 139 residues directly observed in the crystal structure of the NH_2-terminal domain of apoE-3 (Wilson et al., 1991). In addition, the model predicts as random structure 37 residues, which were present in the fragment but not observed in the crystal structure, presumably because of disorder. Two helical segments (1–6, 178–183) have been proposed in the model that occur in disordered regions at the ends of the fragment. The folding of these peripheral helices may be highly dependent on their local environment, requiring either regions of apoE-3 not included in the fragment or lipid to form stable structures. Thirteen of the 23 differently predicted residues occur at the NH_2-terminus of the observed structure in the H_1 helix, an area that was difficult to predict and possibly environmentally sensitive.

ApoA-I model

The integrated prediction scheme resulted in the secondary structure prediction of 93% of the apoA-I primary sequence. A model for apoA-I is shown in Fig. 7 *B* after a unique assignment of the sequence was carried out in a manner similar to apoE-3.

The NH_2-terminal domain (residues 1–57) of apoA-I contains the most ambiguously defined structures. Random coil was assigned from residue 1–8 because of the extended turn prediction and the lack of any dominant feature in the physicochemical analysis of this region. A short region of helix (9–13) with amphipathic character shown by the FPS is predicted to follow. A region of β-structure was assigned between residues 14–22 because of the strong predictions by the secondary structure methods. Residues 23–33 are predicted as random coil. This region contains residues that are uniquely predicted as turn (23–27) or are poorly defined (28–33). Residues 34–40 exhibited low bulk, high hydrophobicity, and had indications of helical structure in only the least stringent selection criteria of the integrated methods. This region was one of the most difficult regions to assign in apoA-I. Since overall helix probability is low

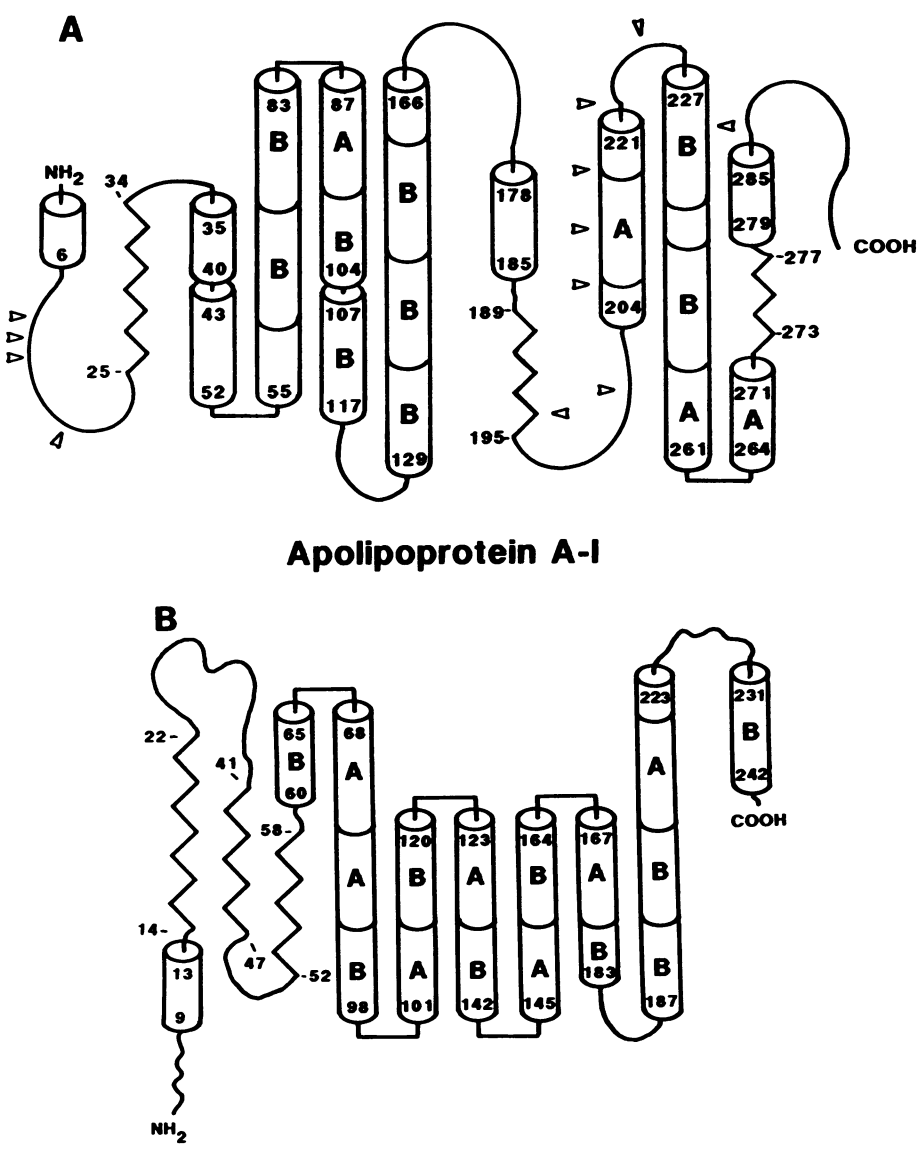

FIGURE 7 ApoE-3 and apoA-1 secondary structure model. A representation of the predicted model for apoE-3 (*A*) and apoA-I (*B*). Helical sections are denoted as cylinders, β structures as ∿∿, and turns as ⊓. Sections of random coil are shown as irregular curves. Assignments to the type A or B consensus sequence are shown on the helical cylinders. ∀ , enzyme cleavage sites reported by Wetterau et al. (1988).

and amphipathic character as evidenced by the FPS is weak, this region has been assigned as random coil.

Residues 41–47 were assigned to β-structure since a significant maximum is observed in the β region of the FPS, and β-structure is uniquely predicted for several residues. A turn follows this region at residues 48–51. The final region assigned to β-structure (52–58) was assigned because of the overall sheet probability and to complete a minimum segment suggested around residue 52. This region was also the best candidate among the undefined segments capable of forming β-structure to meet the overall secondary structure constraint.

The central region of apoA-I (60–183) is comprised solely of α-helices that exhibit varying amphipathic character, as shown by the maxima in the FPS, punctuated by turns. All helices in this region with the exception of the helix between residues 145–164 were uniquely predicted by all methods. The turns in this region were well defined compared with the turns in the central domain of apoE-3, due in part to the greater frequency of proline in the first position of the A-type repeat. Turns were well predicted by the integrated assignment criteria at 3 B/A repeat junctions (99–100, 143–144, 165–166). The turns assigned at B/A junctions (66–67, 121–122) in

which helix was indicated in the final stages of the integrated assignment are less well defined. There is a moderate turn frequency for the residues; however, the FPS does not exhibit a break in the helical region at these positions.

The region between residues 145 and 164 contains an area of low bulk, low helical probability, and lacks significant maxima in the FPS. The integrated prediction methods failed to define this region and indicated helix at each end in only the least stringent protocols. An anomaly appeared to exist in this region between these data and the homology data since this region contains a prototypical A/B consensus sequence repeat with the highest A unit homology (82%) and a significant B unit homology (55%). Modeling of this region in several conformations indicated the clustering of hydrophobic residues on a small helical face. Two Ala residues would be positioned on the hydrophilic face of the helix. The apparent disruption of the FPS is caused by the hydrophobic character of Ala in the scales employed. However, when Ala is considered a neutral residue, an amphipathic helical structure would be indicated in this region. These features are consistent with the consensus sequence. This region is assigned helix as a result of these features.

Residues 184 and 186 are undefined by the integrated protocol, exhibit low bulk due to the pair of glycine residues, are very hydrophilic, and exhibit no significant maximum in the FPS. Thus, these residues are assigned as random coil.

The carboxyl region of apoA-I is predicted as helix from residue 187–223. This region was uniquely defined by the integrated methods with the exception of a short undefined segment (199–202). This segment appears at the transition between a relatively nonamphipathic structure and a strongly amphipathic region containing the highest maximum observed in the FPS. The undefined residues are tentatively assigned helix since there was no indication of a turn at this site. Residues 224–230 exhibit a high hydrophobicity but show no dominant structure potential with both helical character and β-structure evident in this region. This region was assigned as random coil since no clear amphipathic character was evident in the FPS to suggest an ordered structure. Finally, a short amphipathic helix extends from residue 231 to the COOH-terminus.

The composition of the predicted model of apoA-I is 168 residues helix (69%), 23 residues β-structure (9%), and 52 residues assigned to either turn or random coil (21%). These values compare favorably with the CD derived constraints of 70% helix, 10% sheet, and 20% turn/random.

This model for the structure of apoA-I on the surface of a discoidal particle has some significant differences to those presented by others (Segrest et al., 1974; Andrews et al., 1976; Boguski et al., 1986; Li et al., 1988; Brasseur et al., 1990). The model agrees with the previous models in the central region of the molecule from residues 99–186. The longer helices predicted from residues 68–99 and 187–219 have been assigned previously to 22-mer helices and random coil (Marcel et al., 1991). It is interesting to note that a portion of the helix that begins at residue 187 does not exhibit amphipathic character.

The model suggests that there are six major helical segments in the apoA-I molecule, in agreement with the minimal number of helices per apoA-I molecule suggested by Jonas et al. (1989). The punctuation of helices in the central region would allow the exchange of a pair of helices from the solvent to the surface of a lipid particle as first suggested by Brouillette et al. (1984).

These helices, which are composed of multiple A/B tandem repeats, may fold back on each other in an antiparallel fashion. The segregation of charge longitudinally along the helix, clearly demonstrated in the consensus sequence model, suggests that oppositely charged regions may interact when the helices are aligned in this motif. These charge pairing interactions between the helices may stabilize the tertiary packing of this region of apoA-I.

The hydrophilic region around residue 38 may represent a flexible interdomain region. This may imply that the first 37 amino acids of the sequence may represent a single folding domain. However, the antibody binding studies of Marcel et al. (1991) suggest that long-range interactions may occur between the amino terminus and other regions of the molecule. The location of the antibody binding sites identified by this group seem to occur primarily at regions in the model predicted to be in a β-sheet or turn conformation, with the exception of the B type repeat helices located around residues 62, 115, 180, and 204.

General

Analysis of the hydrophobicity patterns has yielded many interesting results for this class of proteins. The concept of an amphipathic helix suggested by Segrest (1977) to be responsible for lipid binding has been extended to several other classes of helix and has led to a classification of helical families (Segrest et al., 1990). The work of De Loof et al. (1986) and our own work (Cladaras et al., 1986) using hydrophobicity and hydrophobic moment analyses has suggested the location of the B/E receptor recognition site in apoE-3 and apoB-100, respectively.

The hydrophobic repeat frequency of α-helical segments suggested by the FPS analysis varies from the ideal value of 100° with significant maxima observed between 80 and 120°. This feature is more evident in the analysis of the apoE-3 sequence than in the apoA-I sequence analysis. The exact repeat frequency may be important in the structure and function of a particular segment, allowing, for example, a more precise alignment of the helical face. Strong hydrophobic repeats were apparent also in the

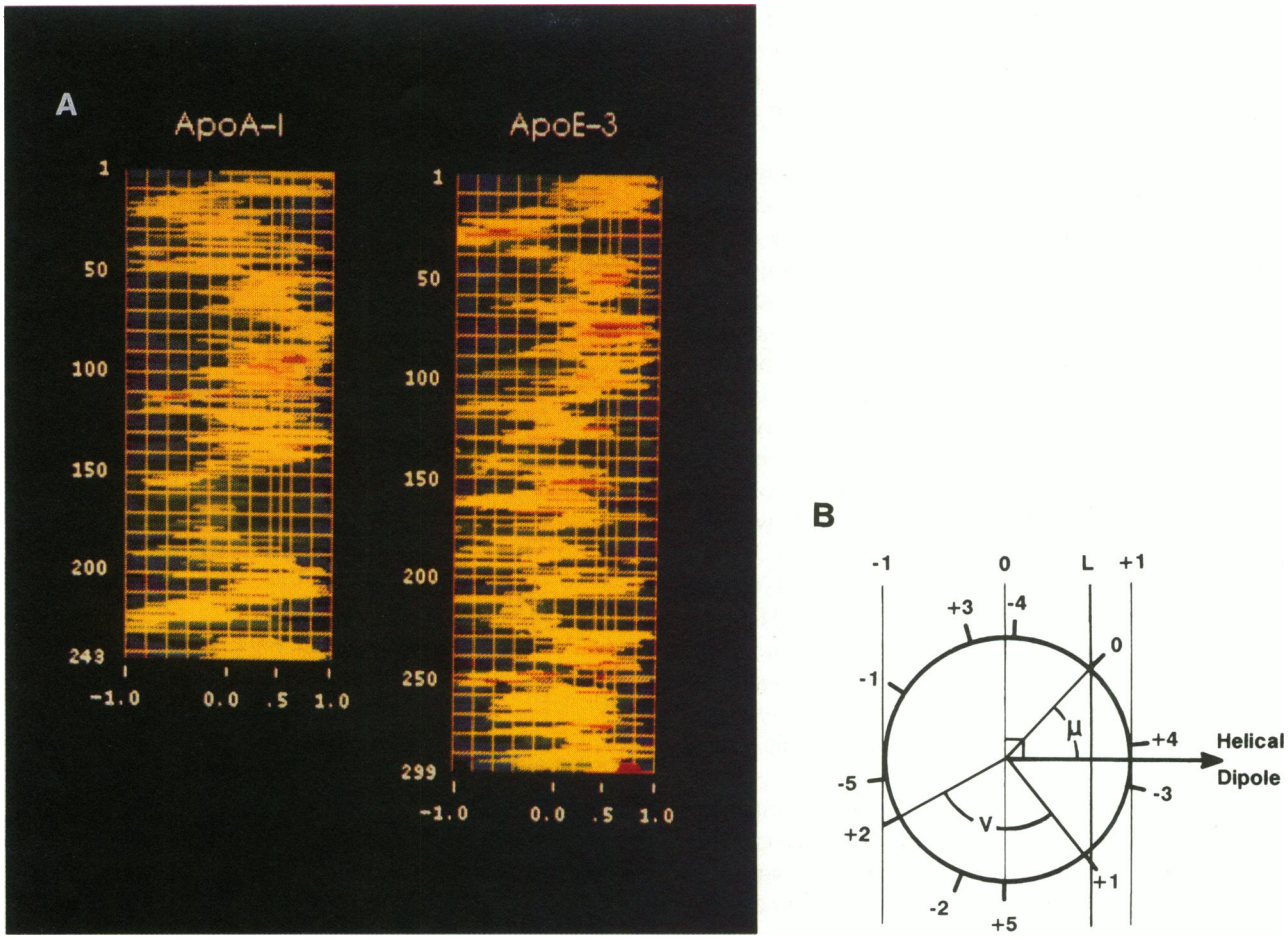

FIGURE 8 Amphipathic helical level calculations for apoE-3 and apoA-I. *A* represents the sum the amino acid hydropathies about a plane of sign inversion in apoE-3 and apoA-I. Each vertical position represents to the center of an 11 amino acid window in the sequence sampled by the calculation. The horizontal axis is the position of a plane of sign inversion placed parallel to the helical axis and perpendicular to the hydrophobic vector of the helix. Each horizontal position represents the sampling of an interfacial level at that position. Grid lines are placed on this axis at 0.20 intervals of the helical radius, with an additional line shown at a 0.5 radial position equivalent to a 120° hydrophobic face. The intensity at each point represents the sum of the hydropathies in each window of the sequence, normalized over all plane positions for that window. Images are displayed with the program SPIDER using the UT1 color scale. *B* shows an illustration of the helical parameters l, v, and μ (defined in the methods) that are involved in this calculation. The helix orientation corresponds to the images with the hydrophobic side shown toward the right.

FPS analysis between 120 and 180°. Other maxima at 150–160° do not correspond to characterized secondary structures. However, it is interesting to speculate that if such a structure is formed at these regions, what physical role it plays in the protein.

To further investigate the relationship of the hydrophobic/hydrophilic faces of the helices within the apolipoproteins, the AHL calculation was developed in the CAPPS program. Fig. 8 contains the results of this calculation on apoA-I and apoE-3. It should be noted that this calculation enforces a helical geometry (with the repeat frequency determined by the maximum value of the FPS in the 80–120° range) on the window of amino acids being studied and therefore is meaningless for those regions of the protein not in a helix. However, variation of the helical geometry is implicit in this analysis.

The calculation of the amphipathic helical interface level along with the power spectra also may reveal information on the finer structural details of the helices that may be responsible for LCAT activity of apoA-I. Fukushima et al. (1980) have proposed that the depth of helical penetration into a lipid monolayer may be an important factor in the activation of LCAT. A maximum activation was postulated by the authors to occur when one-third of the helical surface is hydrophobic, with deeper penetrations into the monolayer surface viewed as disruptive to the structure. Extensive regions within apoA-I are predicted by the AHL analysis to have this interfacial position. Interestingly, apoE-3, which is not a strong activator of LCAT contains only several small regions of helices predicted to penetrate at this depth. This analysis also allows identification of regions that

may penetrate more deeply into the surface of a particle and therefore, could be a measure of the ability of a region to exchange.

The proposed secondary structure assignments enable us to begin to speculate on the possible folding patterns of apoA-I and apoE-3. Examination of the physical properties of the sequences also allow us to propose functional assignments for several of the regions. As an example, the region in apoA-I between 224 and 230 may penetrate deeper into the particle surface and serve as an anchor for the COOH-terminal domain of the molecule.

This class of protein can exist in at least three states: free in plasma, bound to the surface of a nascent HDL disk, and bound at the surface of spherical lipoprotein particles. The models presented were developed using constraints derived from complexes of DMPC and each protein that have been shown to form structures analogous to the HDL disk. Therefore, it should be stressed that these models apply to the proteins as they would appear on the surface of a disk. Other approaches using antibody binding and peptide fragment studies will be invaluable in determining the finer structural features of the proteins. As data from these studies become available, refinements will have to be made to these models.

Although this approach has been tailored specifically to the apolipoproteins in the respect that the first structures assigned are amphipathic in nature, sequences are assigned also at the same constraint levels using average secondary structure probabilities. This later criterion has no inherent amphipathic requirement. The sequential assignment steps are independent; thus, changes in the structure assigned by different criteria are apparent and can be resolved in a final assignment step. The final assignment of ambiguously defined regions provides for the differential weighting of the amphipathic criteria. However, even in the case of globular or membrane spanning proteins, the generation of significant amphipathic potential indicative of a particular structural element would not expect to be preserved evolutionally unless it imparts a structural or functional role.

We thank Dr. Mary Walsh for her assistance with the circular dichroism studies and editorial comments. We also thank Dr. Martha Teeter (Boston College) and Dr. Marc Whitlow (Enzon Inc.) for providing the SEQ program and for help during the initial phases of this study. Finally, we acknowledge Anne Plunkett for the final preparation of this manuscript.

This work was supported by grants from the National Institutes of Health research grant HL-26335 and training grant HL-07291.

Received for publication 19 September 1991 and in final form 13 May 1992.

REFERENCES

Anantharamaiah, G. M., Y. V. Venkatachalapathi, C. G. Brouillette, and J. P. Segrest. 1990. Use of synthetic peptide analogues to localize lecithin: cholesterol acyltransferase activating domain in apolipoprotein A-I. *Arteriosclerosis.* 10:95–105.

Andrews, A. L., D. Atkinson, M. D. Barratt, E. G. Finer, H. Hauser, R. Henry, R. B. Leslie, N. L. Owens, M. C. Phillips, and R. N. Robertson. 1976. Interaction of apoprotein from porcine high-density lipoprotein with dimyristoyl lecithin. Part II. Nature of lipid-protein interaction. *Eur. J. Biochem.* 64:549–563.

Argos, P., M. Hanei, J. M. Wilson, and W. N. Kelly. 1983. A possible nucleotide-binding domain in the tertiary fold of phosphoribosyltransferases. *J. Biol. Chem.* 258:6450–6457.

Atkinson, D., and D. M. Small. 1986. Recombinant lipoproteins: implications for structure and assembly of native lipoproteins. *Annu. Rev. Biophys. Biophys. Chem.* 15:403–456.

Biou, V., J. F. Gilbrat, J. M. Levin, B. Robson, and J. Garnier. 1988. Secondary structure prediction: combination of three different methods. *Protein Eng.* 2:185–191.

Blundell, T. L., and L. N. Johnson. 1976. *Protein Crystallography.* Academic Press, London. 565 pp.

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

Bordo, D., and P. Argos. 1990. Evolution of protein cores: constraints in point mutations as observed in globin tertiary structures. *J. Mol. Biol.* 211:975–988.

Bordo, D., and P. Argos. 1991. Suggestions for safe residue substitutions in site-directed mutagenesis. *J. Mol. Biol.* 217:721–729.

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

Breiter, D. R., M. R. Kanost, M. M. Benning, G. Wesenberg, J. H. Law, M. A. Wells, I. Rayment, and H. M. Holden. 1991. Molecular structure of an apolipoprotein determined at 2.5Å resolution. *Biochemistry.* 30:603–608.

Brewer, H. B., Jr., T. Fairwell, A. LaRue, R. Ronan, A. Houser, and T. J. Bronzert. 1978. The amino acid sequence of human apoA-I, an apolipoprotein isolated from high density lipoproteins. *Biochem. Biophys. Res. Commun.* 80:623–630.

Brouillette, C. G., J. L. Jones, T. C. Ng, H. Kercret, B. H. Chung, and J. P. Segrest. 1984. Structural studies of apolipoprotein A-I/phosphatidylcholine recombinants by high-field proton NMR, non-denaturing gradient gel electrophoresis, and electron microscopy. *Biochemistry.* 23:359–367.

Chou, P. Y., and G. D. Fasman. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* 47:45–148.

Cladaras, C., M. Hadzopoulou-Cladaras, R. T. Nolte, D. Atkinson, and V. I. Zannis. 1986. The complete sequence and structural analysis of human apolipoprotein B-100: relationship between apoB-100 and apoB-48 Forms. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:3495–3507.

Cornette, J. L., K. B. Cease, H. Margalit, J. L. Spouge, J. A. Berzofsky, and C. DeLisi. 1987. Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *J. Mol. Biol.* 195:659–685.

Dayhoff, M. O., W. C. Barker, and L. T. Hunt. 1983. Establishing homologies in protein sequences. *Methods Enzymol.* 91:524–545.

De Loof, H., M. Rosseneu, R. Brasseur, and J. M. Ruysschaert. 1986. Use of hydrophobicity profiles to predict receptor binding domains on apolipoprotein E and the low density lipoprotein apolipoprotein B-E receptor. *Proc. Natl. Acad. Sci. USA.* 83:2295–2299.

De Loof, H., M. Rosseneu, R. Brasseur, and J.-M. Ruysschaert. 1987.

Functional differentiation of amphiphilic helices of the apolipoproteins by hydrophobic moment analysis. *Biochim. Biophys. Acta.* 911:45–52.

Eisenberg, D. 1984. Three-dimensional structure of membrane and surface proteins. *Annu. Rev. Biochem.* 53:595–623.

Eisenberg, D., R. M. Weiss, and T. C. Terwilliger. 1982. The helical hydrophobic moment: a measure of the amphiphilicity of a helix. *Nature (Lond.).* 299:371–374.

Eisenberg, D., E. Schwartz, M. Komaromy, and R. Wall. 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* 179:125–142.

Fielding, C. J., V. G. Shore, and P. E. Fielding. 1972. A protein cofactor of lecithin:cholesterol acyltransferase. *Biochem. Biophys. Res. Commun.* 46:1493–1498.

Finer-Moore, J., and R. M. Stroud. 1984. Amphipathic analysis and possible formation of the ion channel in an acetylcholine receptor. *Proc. Natl. Acad. Sci. USA.* 81:155–159.

Frank, J., B. Shimkin, and H. Dowse. 1981. SPIDER—A molecular software system for electron image processing. *Ultramicroscopy.* 6:343–358.

Fukushima, D., S. Yokoyama, D. J. Kroon, F. J. Kezdy, and E. T. Kaiser. 1980. Chain length-function correlation of amphipathic peptides. *J. Biol. Chem.* 255:10651–10657.

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

Gascuel, O., and J. L. Golmard. 1988. A simple method for predicting the secondary structure of globular proteins: implications and accuracy. *CABIOS.* 4:357–365.

Gilbrat, J.-F., J. Garnier, and B. Robson. 1987. Further developments of protein secondary structure prediction using information theory. *J. Mol. Biol.* 198:425–443.

Greenfield, N., and G. D. Fasman. 1969. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry.* 8:4108–4116.

Hopp, T. P., and K. R. Woods. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. USA.* 78:3824–3828.

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

Jones, D. D. 1975. Amino acid properties and side-chain orientation in proteins: a cross correlation approach. *J. Theor. Biol.* 50:167–183.

Jones, T. A. 1982. FRODO: a graphics fitting program for macromolecules. *In* Computational Crystallography. D. Sayre, editor. Clarendon Press, Oxford. 303–317.

Karathanasis, S. K., V. I. Zannis, and J. L. Breslow. 1983. Isolation and characterization of the human apolipoprotein A-I gene. *Proc. Natl. Acad. Sci. USA.* 80:6147–6151.

Karathanasis, S. K., I. Yunis, and V. I. Zannis. 1986. Structure evolution, and tissue-specific synthesis of human apolipoprotein AIV. *Biochemistry.* 25:3962–3970.

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

Levin, J. M., B. Robson, and J. Garnier. 1986. An algorithm for secondary structure determination in proteins based on sequence similarity. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 205:303–308.

Levitt, M. 1978. Conformational preferences of amino acids in globular proteins. *Biochemistry.* 17:4277–4285.

Li, W.-H., M. Tanimura, C.-C. Lou, S. Datta, and L. Chan. 1988. The apolipoprotein multigene family: biosynthesis, structure-function relationships, and evolution. *J. Lipid Res.* 29:245–271.

Lou, C.-C., W.-H. Li, M. N. Moore, and L. Chan. 1986. Structure and evolution of the apolipoprotein multigene family. *J. Mol. Biol.* 187:325–340.

Lux, S. E., R. Hirz, R. I. Shrager, and A. M. Gotto. 1972. The influence of lipid on the conformation of human plasma high density apolipoproteins. *J. Biol. Chem.* 247:2598–2606.

Mahley, R. W. 1988. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science (Wash. DC).* 240:622–630.

Mahley, R. W., D. Y. Hui, T. L. Innerarity, and K. W. Weisgraber. 1981. Two independent lipoprotein receptors on hepatic membranes of dog, swine, and man. ApoB,E and apo-E receptors. *J. Clin. Invest.* 68:1197–1206.

Mahley, R. W., T. L. Innerarity, S. C. Rall, Jr., and K. H. Weisgraber. 1984. Plasma lipoproteins: apolipoprotein structure and function. *J. Lipid Res.* 25:1277–1294.

Mao, D., and B. A. Wallace. 1984. Differential light scattering and absorption flattening optical effects are minimal in the circular dichroism spectra of small unilamellar vesicles. *Biochemistry.* 23:2667–2673.

Marcel, Y. L., P. R. Provost, H. Koa, E. Raffai, N. V. Dac, J-C. Fruchart, and E. Rassart. 1991. The epitopes of apolipoprotein A-I define distinct structural domains including a mobile middle region. *J. Biol. Chem.* 286:3644–3653.

Markwell, M. A. K., S. M. Haas, L. L. Bieber, and N. E. Tolbert. 1978. A modification of the Lowry procedure to simplify protein determination in membrane and lipoprotein samples. *Anal. Biochem.* 87:206–210.

Matthews, B. W. 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta.* 405:442–451.

McLachlan, A. D. 1977. Repeated helical pattern in apolipoprotein A-I. *Nature (Lond.).* 267:465–466.

Nishikawa, K., and T. Ooi. 1986. Amino acid sequence homology applied to the prediction of protein secondary structures, and joint prediction with existing methods. *Biochim. Biophys. Acta.* 871:45–54.

Prevelige, P., Jr., and G. D. Fasman. 1989. Chou-Fasman prediction of secondary structure of proteins: the Chou-Fasman-Prevelige algorithm. *In* Prediction of Protein Structure and the Principles of Protein Conformation. G. Fasman, editor. Plenum Publishing Corp., New York. 391–416.

Rall, S. C., Jr., K. H. Weisgraber, and R. W. Mahley. 1982. Human apolipoprotein E. The complete amino acid sequence. *J. Biol. Chem.* 257:4171–4178.

Roth, R. I., R. L. Jackson, H. J. Pownall, and A. M. Gotto, Jr. 1977. Interaction of plasma 'arginine-rich' apolipoprotein with dimyristoylphosphatidylcholine. *Biochemistry.* 16:5030–5036.

Scanu, A., J. Toth, C. Edelstein, S. Koga, and E. Stiller. 1969. Fractionation of human serum high density lipoprotein in urea solutions. Evidence for polypeptide heterogeneity. *Biochemistry.* 8:3309–3316.

Schultz, G. E., C. D. Barry, J. Friedman, P. Y. Chou, G. D. Fasman, A. V. Finkelstein, V. I. Lim, O. B. Ptitsyn, E. A. Kabat, T. T. Wu, M. Levitt, B. Robson, and K. Nagano. 1974. Comparison of predicted and experimentally determined secondary structure of adenyl kinase. *Nature (Lond.).* 250:140–142.

Segrest, J. P. 1977. Amphipathic helixes and plasma lipoproteins: thermodynamic and geometric considerations. *Chem. Phys. Lipids.* 18:7–22.

Segrest, J. P., and R. J. Feldmann. 1977. Amphipathic helixes and plasma lipoproteins: a computer study. *Biopolymers.* 16:2053–2065.

Segrest, J. P., R. L. Jackson, J. D. Morrisett, and A. M. Gotto, Jr. 1974. A molecular theory of lipid-protein interactions in the plasma lipoproteins. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 38:247–253.

Segrest, J. P., H. De Loof, J. G. Dohlman, C. G. Brouillette, and G. M. Anantharamaiah. 1990. Amphipathic helix motif: classes and properties. *Proteins.* 8:103–117.

Shore, V. G., R. E. Garcia, A. L. Penn, and B. Shore. 1980. Polyacrylamide gel electrophoresis and isoelectric focusing of plasma apolipoproteins. *In* CRC Handbook of Electrophoresis. Vol. I. Lipoproteins: Basis, Principles and Concepts. L. A. Lewis and J. J. Opplt, editors. CRC Press, Inc., Boca Raton, FL. 103–125.

Sparrow, J. T., and A. M. Gotto, Jr. 1982. Apolipoprotein/lipid interactions: studies with synthetic polypeptides. *CRC Crit. Rev. Biochem.* 13:87–107.

Steinmetz, A., and G. Utermann. 1985. Activation of lecithin:cholesterol acyltransferase by human apolipoprotein A-IV. *J. Biol. Chem.* 260:2258–2264.

Teeter, M. M., and M. Whitlow. 1988. Test of circular dichroism (CD) methods for crambin and CD-assisted secondary structure prediction of its homologous toxins. *Proteins.* 4:262–273.

Walsh, M. T., and D. Atkinson. 1986. Physical properties of apoprotein B in mixed micelles with sodium deoxycholate and in a vesicle with dimyristoyl phosphatidylcholine. *J. Lipid Res.* 27:316–325.

Weisgraber, K. H., T. L. Innerarity, K. J. Harder, R. W. Mahley, R. W. Milne, Y. L. Marcel, and J. T. Sparrow. 1983. The receptor binding domain of apolipoprotein E. Monoclonal antibody inhibition of binding. *J. Biol. Chem.* 258:12348–12354.

Wetterau, J. R., and A. Jonas. 1982. Effect of dipalmitoylphosphatidylcholine vesicle curvature on the reaction with human apolipoprotein A-I. *J. Biol. Chem.* 257:10961–10966.

Wetterau, J. R., L. P. Aggerbeck, S. C. Rall, Jr., and K. H. Weisgraber. 1988. Human apolipoprotein E3 in aqueous solution. Evidence for two structural domains. *J. Biol. Chem.* 263:6240–6248.

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