Covalent Structure of Apolipoprotein A-II from *Macaca mulatta* Serum High-Density Lipoproteins†

Celina Edelstein, Claudia Noyes, Pamela Keim, Robert L. Heinrikson, Robert E. Fellows, and Angelo M. Scanu*

ABSTRACT: The covalent structure of apolipoprotein A-II, isolated from the serum high-density lipoprotein of a single male Rhesus monkey (*Macaca mulatta*), was determined. The amino acid sequence of this 77-residue polypeptide is: <Glu-Ala-Glu-Glu-Pro^{5}-Ser-Val-Glu-Ser-Leu^{10}-Val-Ser-Gln-Tyr-Phe^{15}-Gln-Thr-Val-Thr-Asp^{20}-Tyr-Gly-Lys-Asp-Leu^{25}-Met-Glu-Lys-Val-Lys^{30}-Ser-Pro-Glu-Leu-Gln^{35}-Ala-Gln-Ala-Lys-Ala^{40}-Tyr-Phe-Glu-Lys-Ser^{45}-Lys-Glu-Gln-Leu-Thr^{50}-Pro-Leu-Val-Lys-Lys^{55}-Ala-Gly-Thr-Asp-Leu^{60}-Val-Asn-Phe-Leu-Ser^{65}-Tyr-Phe-Val-Glu-Leu^{70}-Arg-Thr-Gln-Pro-Ala^{75}-Thr-Gln-COOH. A comparison of this structure to that of the monomeric form of human apolipoprotein A-II reveals a high degree of homology except for six conservative amino acid replacements (positions 3, 6, 40, 53, 59, and 71). Of particular structural significance is the replacement of cysteine by serine in position 6. This explains why Rhesus A-II exists in monomeric form, contrary to the established dimeric nature of the human protein.

The physical and chemical properties of Rhesus monkey (*Macaca mulatta*) serum high-density lipoprotein (HDL)[1] have been shown to be similar but not identical with those of human HDL (Scanu et al., 1973). Studies of the Rhesus HDL apoprotein (Edelstein et al., 1973) revealed that its two major polypeptides, A-I and A-II, exist in a weight ratio of 3:1, respectively, and have chemical and physical parameters very similar to those of the corresponding human products. However, whereas the A-II of Rhesus monkey was found to be a monomer of mol wt 8757, the human protein is a dimer (mol wt 17 410) composed of two identical monomers bonded by a single disulfide bridge (Brewer et al., 1972). Since the primary structure of human A-II protein had been determined by Brewer et al. (1972), it was of special importance, from both the chemical and phylogenetic viewpoints, to determine the amino acid sequence of Rhesus A-II. Such studies are the subject of this report. Preliminary results have appeared elsewhere (Edelstein et al., 1974).

Experimental Section

Materials. Serum was obtained from the blood of a single male Rhesus monkey, age 1–2 years, maintained on a modified Purina Primate Chow diet which was low in fat and cholesterol. Plasmapheresis was conducted at about 2-week intervals and did not affect the animal adversely, as assessed by periodic clinical and laboratory studies. During the whole period of the experiment, the serum cholesterol levels averaged between 100 and 150 mg/100 ml. High-density lipoproteins were isolated by preparative ultracentrifugation as described previously (Scanu et al., 1973) and were shown to be free of low-density lipoproteins and other serum proteins by agarose electrophoresis (Noble, 1968) and immunochemical methods (Scanu et al., 1973). The lipid-free apo-HDL was obtained by ethanol–ether extraction of HDL at −10 °C (Scanu and Edelstein, 1971). Rhesus A-II was isolated by a combination of Sephadex and DEAE chromatography in 6 M urea as reported earlier (Edelstein et al., 1973). The purity of the protein so isolated was ascertained by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate and 8 M urea, determination of COOH- and NH_2-terminal residues, compositional analysis, and immunological criteria (Edelstein et al., 1973).

Ethanol, diethyl ether, and *N*-ethylmorpholine were reagent grade and were redistilled immediately before use. Guanidine hydrochloride, Ultra Pure, was obtained from Heico (Delaware Water Gap, Pa.). Urea (Baker, Chicago, Ill.) was recrystallized from ethanol and solutions thereof were deionized on a mixed bed resin (AG-501 X8CD, 20–50 mesh, Bio-Rad Laboratories, Richmond, Calif.). Cyanogen bromide was purchased from Pierce Chemical Co. (Rockford, Ill.). Carboxypeptidases A and B (treated with diisopropyl phosphofluoridate) were the products of Sigma Chemical Co. (St. Louis, Mo.), and trypsin (code TRTPCK-1FA) was obtained from Worthington Biochemicals (Freehold, N.J.). Carboxypeptidase Y was prepared following the procedure of Hayashi et al. (1973).

Pyroglutamyl aminopeptidase (<Glu-peptidase) was isolated from freeze-dried cultures of *Bacillus subtilis* (Sigma Chemical Co., St. Louis, Mo.) by ammonium sulfate precipitation and gel filtration on Sephadex G-150 (Fellows, R., Mudge, R., and Graves, J. C., manuscript in preparation). This enzyme hydrolyzed synthetic <Glu-Ala (Cyclo Chemical Co., Los Angeles, Calif.) at a rate of 0.25 μmol/min at 40 °C and was free of demonstrable endopeptidase activity.

Amino Acid Analysis. Proteins and peptides were hydro-

† From the Departments of Biochemistry and Medicine, the University of Chicago Pritzker School of Medicine, and the Franklin McLean Memorial Research Institute (operated by the University of Chicago for the United States Energy Research and Development Administration), Chicago, Illinois 60637 (C.E., C.N., P.K., R.L.H., and A.M.S.), and the Department of Physiology, Duke University Medical Center, Durham, North Carolina 27710 (R.E.F.). *Received September 16, 1974.* Supported by Grants HL-08727 and HL-06481 from the United States Public Health Service, Grant A72-6 from the Chicago and Illinois Heart Association, and Grant GB-29098 from the National Science Foundation. A.M.S. is recipient of a U.S. Public Health Service Career Development Award (HL-24,867).

[1] The abbreviations used are: HDL, high-density lipoproteins of *d* 1.063–1.21 g/ml; apo-HDL, delipidated HDL; A-II, apolipoprotein A-II from apo-HDL; <Glu, pyroglutamic acid; <Glu-peptidase, pyroglutamyl aminopeptidase EC 3.4.11.8; Pth, phenylthiohydantoin.

lyzed in 6 N HCl according to the general procedures of Moore and Stein (1963) and the hydrolysates were analyzed on either a Beckman Model 121 amino acid analyzer (Beckman Instruments, Palo Alto, Calif.) or a Durrum D-500 (Durrum Instruments, Palo Alto, Calif.) amino acid analyzer (Sterner et al., 1974). Some of the phenylthiohydantoin (Pth) derivatives obtained during the course of automated Edman degradation were converted to the parent amino acid or an appropriate derivative by hydrolysis in hydroiodic acid as described by Smithies et al. (1971); the products were, in turn, identified by amino acid analysis. Protein determinations were made by the Lowry procedure (Lowry et al., 1951) or by computations from amino acid analysis.

High-Voltage Paper Electrophoresis. High-voltage paper electrophoresis of peptides was performed at pH 6.5 (Bennett, 1967) or at pH 2.0 (0.61 N formic acid) in a Gilson Model D Electrophorator. Samples were applied onto Whatman No. 1 paper, and electrophoresis was carried out at 3000 V for 45 min. Electropherograms were dried, and peptides were visualized under ultraviolet light after the papers had been sprayed with fluorescamine (Udenfriend et al., 1972) or an arginine-sensitive phenanthrenequinone reagent (Yamada and Itano, 1966).

Polyacrylamide Gel Electrophoresis in Sodium Dodecyl Sulfate and in 8 M Urea. Gels were run in sodium dodecyl sulfate according to the method of Weber and Osborn (1969). Urea gels containing 7.5% acrylamide were run as described by Davis (1964) with a modification for the inclusion of 8 M urea (Scanu et al., 1969).

Cyanogen Bromide Cleavage at Methionine. Rhesus A-II protein (10 mg) was dissolved in 3 ml of 70% formic acid. A 500-fold molar excess of cyanogen bromide was added, and the vessel was sealed under nitrogen. After reaction for 48 h at room temperature, an aliquot was removed for amino acid analysis for verification and quantitation of the extent of cleavage based on the appearance of homoserine and homoserine lactone and the disappearance of methionine. Fractionation of the cyanogen bromide fragments was achieved by gel filtration on a column (1.5 × 200 cm) of Sephadex G-50 as described previously (Edelstein et al., 1974) and their purity was ascertained by high-voltage paper electrophoresis, amino acid analysis, and COOH- and NH_2-terminal analyses.

Succinylation. The lysine residues were blocked at pH 8.2 by reaction with succinic anhydride. Rhesus A-II (10 mg) was dissolved in 3 ml of 0.02 M phosphate buffer (pH 8.2) containing 6 M guanidine hydrochloride, and a 60:1 molar ratio of anhydride to lysine residues was added in small portions over a period of 2 h at room temperature. The reaction was carried out in a pH stat (Radiometer, Copenhagen, Denmark) with 1 N NaOH as titrant to maintain a constant pH of 8.2. The modified protein was desalted on a column (2.5 × 50 cm) of Sephadex G-10 equilibrated with 0.01 M *N*-ethylmorpholine acetate buffer (pH 8.2) and the extent of succinylation was determined with 2,4,6-trinitrobenzenesulfonic acid (Fields, 1972); the reaction was followed spectrophotometrically at 420 nm.

Trypsin Digestion. Rhesus A-II and cyanogen bromide fragments derived therefrom were digested with 1% by weight of trypsin in 0.1 M *N*-ethylmorpholine acetate buffer (pH 8.0). Digestions were carried out for 1 h at 37 °C, and reactions were stopped by the addition of urea to a concentration of 8 M, or by acidification to pH 2.0.

Isolation of the Cleavage Products Obtained in the Digestion of Native and Succinylated A-II with Trypsin. The tryptic digest of succinylated A-II was fractionated at 8 °C on a column (1.5 × 200 cm) of Sephadex G-50 (superfine) in 0.05 M Tris (pH 8.2) containing 8 M urea. Effluent fractions of 3 ml were monitored for peptide content by measuring the absorption at 280 nm and by reacting aliquots of each fraction which had been spotted on Whatman No. 1 paper with fluorescamine (Udenfriend et al., 1972).

The tryptic digest of native A-II was fractionated at room temperature on a column (1.5 × 200 cm) of Sephadex G-25 (superfine) in 0.1 M ammonium bicarbonate (pH 8.0). Further purification of certain fractions was achieved by ion exchange chromatography on a column (1.5 × 60 cm) of DEAE-cellulose developed at pH 8.2 with a linear gradient of increasing ammonium bicarbonate concentration from 0.01 to 0.1 M. Fractions of 1.5 ml, collected at a flow rate of 15 ml/h, were monitored spectrophotometrically at 225 and 280 nm for detection of peptides.

Sequence Analysis. Carboxyl-terminal sequences were derived by the stoichiometry and kinetics of amino acid liberation during digestion with carboxypeptidases A, B, or Y. Digestion with carboxypeptidase A or B was carried out at 25 °C (50:1, protein:enzyme molar ratio) in 0.2 M *N*-ethylmorpholine acetate buffer (pH 8.5). Aliquots were removed at timed intervals and the reactions were terminated by quick freezing in a dry ice–acetone bath, followed by lyophilization. The released amino acids were analyzed with a Beckman 121 amino acid analyzer equipped with a single column lithium chloride buffer system capable of resolving glutamine and asparagine (Beckman Bulletin A-TB-044). Hydrolysis with carboxypeptidase Y was performed in a similar fashion except that the buffer was 0.1 M pyridine acetate (pH 5.5). To ensure the release of acidic amino acids in the COOH-terminal region, certain tryptic peptides were digested initially with carboxypeptidase B followed by lowering of the pH to 5.4 with acetic acid and digestion with carboxypeptidase A at 30 °C.

Protein or peptide samples with blocked amino termini were digested with <Glu-peptidase, and the liberation of pyroglutamic acid (<Glu) was quantitated by gas chromatography as described previously (Edelstein et al., 1974).

Automated Edman degradation (Edman and Begg, 1967) was performed in a Beckman protein–peptide sequencer (Model 890B) equipped with nitrogen flush and undercut reaction cup accessories. Programs employing *N,N*-dimethylallylamine as the coupling buffer were a single-cleavage peptide program (Beckman No. 090872) and a modification thereof employing double cleavage. The Pth-amino acids liberated after each cycle of the degradation were identified as such or as the trimethylsilyl derivatives by gas chromatography (Pisano and Bronzert, 1969) on a Beckman GC-45 unit fitted with 2 mm i.d. × 4 ft glass columns packed with 10% SP400 coated on Chromosorb W (HP), 100–120 mesh. In some cases, identifications were confirmed by thin-layer chromatography (Jeppson and Sjöquist, 1967) and the chromatograms were stained with ninhydrin to produce distinctive colors for certain of the Pth derivatives (Inagami and Murakami, 1972). Hydrolytic back conversion in hydroiodic acid (Smithies et al., 1971), followed by amino acid analysis, was also employed in some cases to confirm the identification and to provide quantitation of certain residues.

Results

The amino acid composition of HDL Rhesus A-II pre-

Table I: Amino Acid Composition of Rhesus A-II and Fragments Derived therefrom by Cleavage with Cyanogen Bromide.

Amino Acid	Residues/mol[c]		
	Apo A-II[a]	CNBr-I[b]	CNBr-II[b]
Asp + Asn	4.3 ± 0.20 (4)	2.0 ± 0.05 (2)	2.2 ± 0.55 (2)
Thr[d]	5.8 ± 0.23 (6)	2.1 ± 0.05 (2)	4.1 ± 0.58 (4)
Ser[d]	5.7 ± 0.28 (6)	3.0 ± 0.06 (3)	3.0 ± 0.32 (3)
Glu + Gln	16.4 ± 0.95 (16)	6.1 ± 0.22 (6)	10.1 ± 0.68 (10)
Pro	4.0 ± 0.21 (4)	0.8 ± 0.09 (1)	2.9 ± 0.25 (3)
Gly	2.2 ± 0.16 (2)	1.0 ± 0.23 (1)	1.1 ± 0.01 (1)
Ala	6.0 (6)	1.0 (1)	5.0 (5)
Val	6.8 ± 0.23 (7)	3.1 ± 0.09 (3)	4.4 ± 0.65 (4)
Met	0.8 ± 0.09 (1)	0 (0)	0 (0)
Leu	8.1 ± 0.23 (8)	1.7 ± 0.07 (2)	5.9 ± 0.13 (6)
Tyr	3.8 ± 0.35 (4)	2.0 ± 0.09 (2)	2.1 ± 0.52 (2)
Phe	4.0 ± 0.16 (4)	0.8 ± 0.06 (1)	2.7 ± 0.11 (3)
Lys	7.9 ± 0.35 (8)	1.1 ± 0.27 (1)	7.2 ± 0.05 (7)
Arg	1.0 ± 0.08 (1)	0 (0)	1.0 ± 0.10 (1)
Hse + lactone	0 (0)	0.9 ± 0.01 (1)	0 (0)
Total no. of residues	77	26	51
NH_2-terminal residue	<Glu	<Glu	Glu
COOH-terminal residue	Gln	Hse	Gln

[a] Residue numbers ± standard deviation of the mean were calculated on the basis of 12 analyses. [b] Residue numbers ± standard deviation of the mean were calculated on the basis of four analyses. [c] Numbers in parentheses calculated to nearest integer and based on alanine = 6, 1, and 5 residues per mol of A-II, CNBr-I, and CNBr-II, respectively. [d] Corrected for losses during hydrolysis as assessed by hydrolyses for 18, 24, 48, and 72 h. These losses were 5% for threonine and 10% for serine.

Table II: Amino Acid Composition of Tryptic Fragments from Succinylated Rhesus A-II.

Amino Acid	Residues/mol[c]	
	S-I[a]	S-II[b]
Asp + Asn	4.0 ± 0.10 (4)	
Thr[d]	3.9 ± 0.10 (4)	2.0 ± 0.06 (2)
Ser[d]	5.8 ± 0.05 (6)	
Glu + Gln	14.2 ± 0.74 (14)	2.3 ± 0.03 (2)
Pro	2.7 ± 0.02 (3)	1.2 ± 0.10 (1)
Gly	2.1 ± 0.03 (2)	
Ala	5.0 (5)	1.0 (1)
Val	6.7 ± 0.82 (7)	
Met	0.9 ± 0.10 (1)	
Leu	8.1 ± 0.79 (8)	
Tyr	3.9 ± 0.55 (4)	
Phe	3.8 ± 0.18 (4)	
Lys	7.9 ± 0.44 (8)	
Arg	0.9 ± 0.06 (1)	
Total no. of residues	71	6
NH_2-terminal residue	<Glu	Thr
COOH-terminal residue	Arg	Gln

[a] Based on 5 mol of alanine; number in parentheses indicates moles of amino acid to the nearest whole integer. [b] Based on 1 mol of alanine per mol of peptide. [c] Values and their standard deviations based on four analyses. [d] Adjusted for hydrolytic losses, based on hydrolyses for 18, 24, 48, and 72 h. These losses were 5% for threonine and 10% for serine.

sented in Table I dictated the general strategy for determination of the covalent structure of this 77-residue polypeptide. The absence of cysteine or cystine, tryptophan, isoleucine, and histidine should be noted as well as the fact that the protein contains single residues of methionine and arginine. The existence of these two distinct sites for specific chemical and enzymic cleavage of Rhesus A-II greatly facilitated the approach to sequence analysis by automated Edman degradation. In an earlier account we reported that Rhesus apo A-II had 78 residues (Edelstein et al., 1973). A careful statistical analysis of several amino acid determinations together with the present sequence data now establishes that the protein contains a total of 77 residues and that in the previous analysis glutamic acid was overestimated by one residue.

Earlier reports from this laboratory (Edelstein et al., 1973, 1974) indicated that the amino terminus of Rhesus A-II is blocked by <Glu. Thus, to make the protein amenable to automated Edman degradation, a sample of A-II (0.60 μmol) was treated with <Glu-peptidase. After the blocking residue was removed in this way and identified as previously described (Edelstein et al., 1974) the residual polypeptide (0.52 μmol) was subjected to 43 stages of automated Edman degradation. The sequence of the first 44 residues thus obtained[2] contains quantitative data generated during a typical analysis. The identifications were based upon quantitative gas chromatographic analysis[2] (Pisano and Bronzert, 1969) of the Pth amino acids and amino acid analysis of the products resulting from them by hydroiodic acid hydrolysis (Smithies et al., 1971) (cf. Experimental Section). The repetitive yield calculated from the recoveries of Pth-Val-7 and Pth-Val-29 was 93%, and the initial yield of Pth-Ala-2 was 82%. It is noteworthy that the single half-cystine residue in position 6 of the human A-II protein, which accounts for its existence as a dimer (Brewer et al., 1972), is replaced by serine in the Rhesus sequence.[2] Further differences are the substitution of Lys-3 and Ser-40 in human A-II by glutamic acid and alanine, respectively, in Rhesus A-II.[2]

The location of the single methionine in A-II at position 26 was of strategic importance in the covalent structural analysis by automated Edman degradation. Cleavage of the intact protein by cyanogen bromide gave two fragments the properties of which were reported previously (Edelstein et al., 1974). As expected, the NH_2-terminal peptide of 26 residues (CNBr-I) was blocked with pyroglutamic acid and terminated in homoserine lactone. The compositional analysis of CNBr-I given in Table I is in accord with the sequence determination of the first 44 residues.[2] In agreement with the earlier findings (Edelstein et al., 1974), the COOH terminal fragment of 51 residues (CNBr-II) had glutamic acid at the amino terminus and glutamine at the COOH terminus and did not contain homoserine. The amino acid composition of this fragment was previously overestimated by one glutamic acid residue. The sequence of residues 27–75 in the intact A-II chain was obtained by automated Edman degradation of CNBr-II (0.65 μmol).[2] This analysis overlapped and confirmed 18 of the 44 residues sequenced previously.[2] The repetitive yield based upon recoveries of Pth-Val at cycles 3 and 27 was 95%. In addition to the

[2] The results of these studies, contained in four tables and one figure, were submitted in detail for examination by the reviewers and will be available as supplementary material; see paragraph at end of paper.

Table III: Analysis of the Tryptic Peptides of Rhesus A-II.

Amino Acid	T-1	T-I′	T-II	T-III	T-IV	T-V	T-VI	T-VII′	T-VII	T-VIII	T-IX
Asp	1.1 (1)	1.8 (2)	0.9 (1)						1.9 (2)	2.0 (2)	
Thr	1.8 (2)	1.7 (2)					0.9 (1)		0.9 (1)	0.9 (1)	2.1 (2)
Ser	2.8 (3)	2.8 (3)			1.0 (1)		1.0 (1)		1.1 (1)	1.3 (1)	
Glu	5.8 (6)	6.7 (7)	0.9 (1)		2.7 (3)	1.0 (1)	2.3 (2)		1.0 (1)	1.3 (1)	2.1 (2)
Pro	0.9 (1)	0.8 (1)			0.9 (1)		1.0 (1)				1.0 (1)
Gly	1.0 (1)	1.2 (1)							1.1 (1)	1.4 (1)	
Ala	1.0 (1)	1.0 (1)			1.7 (2)	1.0 (1)			1.0 (1)	1.0 (1)	1.0 (1)
Val	2.8 (3)	2.8 (3)		1.0 (1)			1.0 (1)		1.7 (2)	1.9 (2)	
Met		0.8 (1)	0.8 (1)								
Leu	1.0 (1)	1.7 (2)	1.0 (1)		1.0 (1)		1.7 (2)		2.9 (3)	2.9 (3)	
Tyr	1.7 (2)	1.8 (2)				1.0 (1)			1.0 (1)	0.8 (1)	
Phe	0.9 (1)	0.8 (1)				1.0 (1)			1.8 (2)	1.8 (2)	
Lys	1.0 (1)	1.7 (2)	0.8 (1)	0.8 (1)	1.0 (1)	1.1 (1)	1.8 (2)	1.0 (1)	0.9 (1)		
Arg									0.9 (1)	0.8 (1)	
Total no. residues	23	28	5	2	9	5	10	1	17	16	6
NH_2-terminal	<Glu	<Glu	Asp	Val	Ser	Ala	Ser	ND	Lys	Ala	Thr
COOH-terminal	Lys	Lys	Lys	Lys	Lys	Lys	Lys	ND	Arg[a]	Arg	Gln[b]
Method of purification	Sephadex G-25	Sephadex G-25	DEAE-cellulose	Sephadex G-25	DEAE-cellulose	Sephadex G-25	Sephadex G-25	Sephadex G-25	DEAE-cellulose	DEAE-cellulose	DEAE-cellulose
% yield	40	45	20	30	45	80	80	10	40	60	90
Residues in sequence	1-23	1-28	24-28	29-30	31-39	40-44	45-54	55	55-71	56-71	72-77
Peak	B	A	DE-5	1	DE-1	G	C	2	DE-4	DE-6	DE-2

[a] Determined with carboxypeptidase B only. [b] Determined with carboxypeptidases A and Y. ND = not determined.

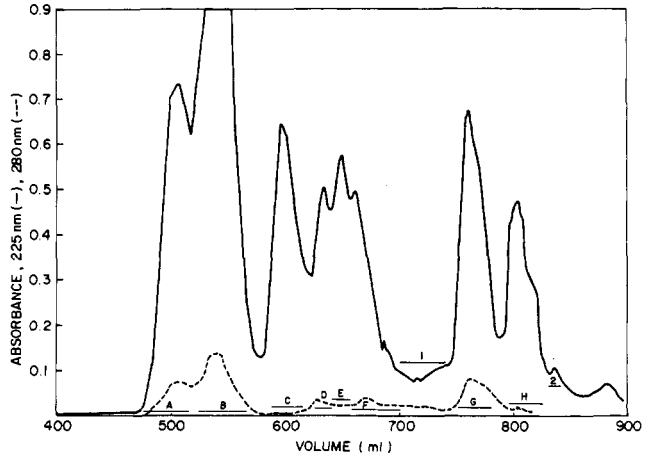

FIGURE 1: Elution profile of Rhesus A-II after digestion with trypsin. Conditions: Sephadex G-25 (superfine) column (1.5 × 200 cm) eluted with 0.1 M ammonium bicarbonate (pH 8.0) at room temperature.

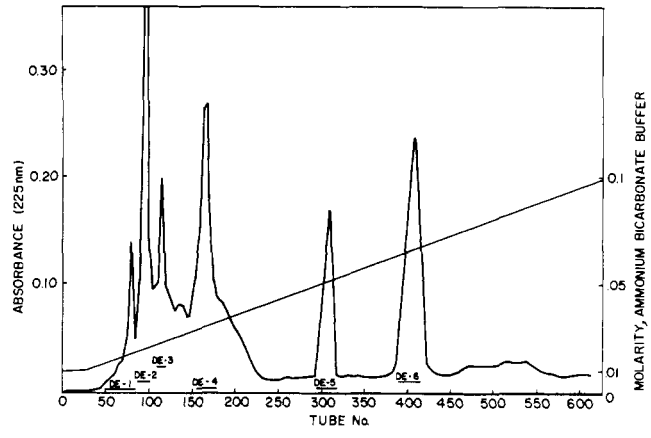

FIGURE 2: DEAE chromatography of a mixture of tryptic peptides D, E, and F, obtained by molecular sieve chromatography as shown in Figure 1 and explained in the text. Conditions: DE-52 cellulose column (1.5 × 60 cm) eluted with a linear gradient of ammonium bicarbonate (0.01 to 0.1 M), pH 8.2; flow rate, 15 ml/h; room temperature.

aforementioned replacements of Lys-3, Cys-6, and Ser-40 in human A-II by glutamic acid, serine, and alanine, respectively, in the monkey protein, differences at three other loci were observed. Residue 53 was identified as valine, residue 59 as aspartic acid, and residue 71 as arginine, rather than isoleucine, glutamic acid, and glycine, respectively, which appear in those positions in human A-II (Brewer et al., 1972).

To confirm and extend the sequence obtained by analysis of the cyanogen bromide fragments presented separately[2] and in Table I, tryptic peptides were isolated from apo A-II and subjected to compositional and sequence analysis. Two types of experiments were conducted. In the first, Rhesus A-II was succinylated to block the lysine residues and the polypeptide was cleaved with trypsin at the single arginine (residue 71). This tryptic digest was separated by gel chromatography on Sephadex G-50 into two fractions (S-I and S-II), the amino acid compositions of which are given in Table II. Fragment S-I contained 71 residues and included the single residue of arginine. From the kinetics of carboxypeptidase Y digestion of S-I,[2] the following COOH-terminal sequence was reconstructed: -Val-Glu-Leu-Arg-COOH. These findings, shown separately,[2] confirmed the assignments made previously by Edman degradation of CNBr-II.[2] Fragment S-II constituted the COOH-terminal hexapeptide in A-II (Table II); Edman degradation of 0.5 μmol of this fragment gave the sequence Thr-Gln-Pro-Ala-Thr-Gln-COOH. Threonine and glutamine were detected by gas–liquid and thin-layer chromatography. The yields for proline and alanine were 90 and 96%, respectively.

COOH-Terminal Analysis of A-II. Digestion of native

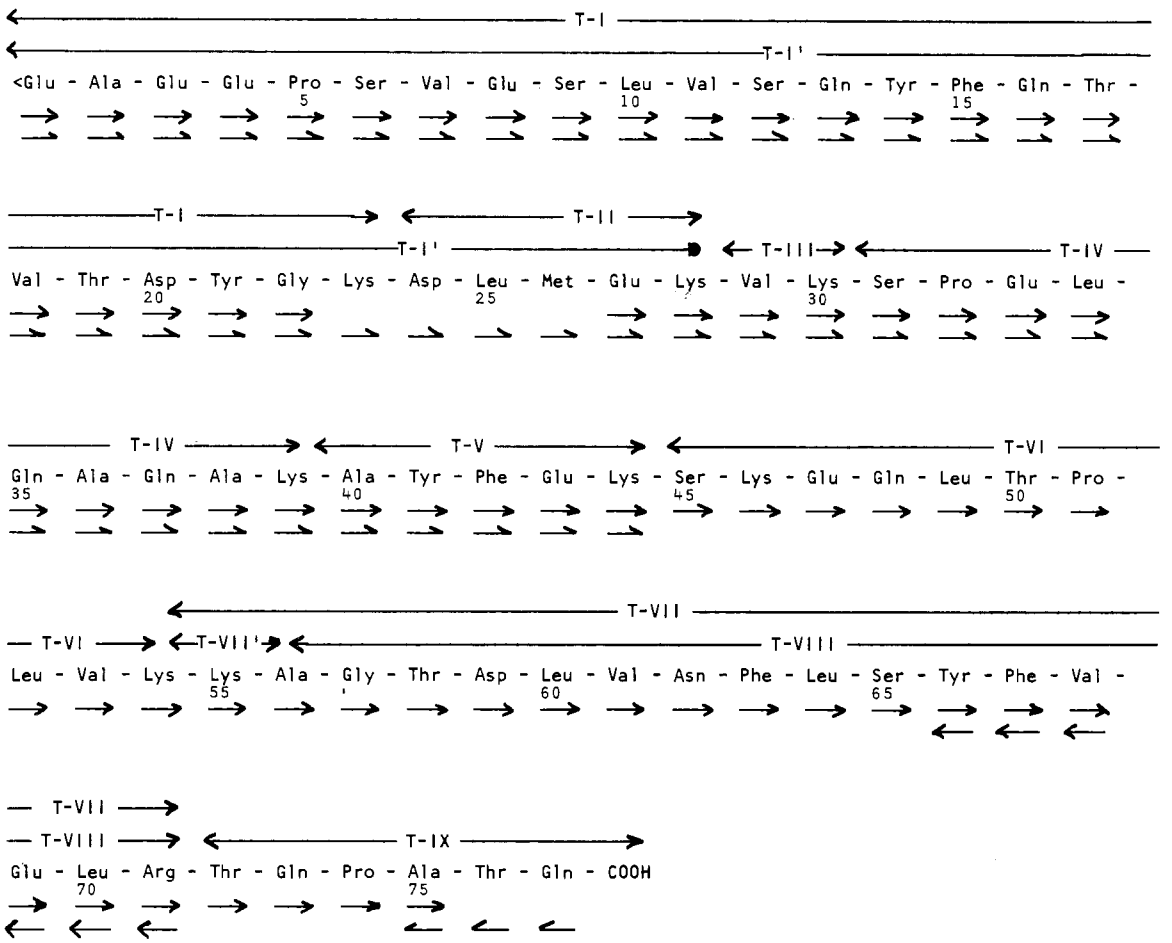

FIGURE 3: Complete covalent structure of Rhesus A-II with a schematic representation of the sequence steps performed in the analysis: (→) automated Edman degradation of <Glu-peptidase-treated Rhesus A-II; (→) automated Edman degradation of CNBr-I treated with <Glu-peptidase, and CNBr-II showing the overlap of 18 residues (27–44); (←) carboxypeptidase Y digestion of tryptic peptide S-I; (←) carboxypeptidase Y digestion of intact Rhesus A-II. Tryptic peptides designated T-I to T-IX were isolated and sequenced by manual and automated Edman degradation as described in the text.

A-II with carboxypeptidase A gave the carboxyl-terminal sequence -Thr-Gln-COOH (Edelstein et al., 1974). In the present study, carboxypeptidase Y digestion gave -Ala-Thr-Gln-COOH as the COOH-terminal sequence.[2] These findings are in accord with the sequence analysis of the COOH-terminal hexapeptide (S-II) mentioned above.

Analysis of Tryptic Fragments. Native A-II (2 μmol) was digested with trypsin and the peptides were separated by gel filtration on a column (2.5 × 200 cm) of Sephadex G-25 (superfine) eluted with 0.1 M ammonium bicarbonate (Figure 1). Peaks A and B were shown by amino acid analysis to be the peptides containing residues 1 to 28 and 1 to 23, respectively. Since the sequence data[2] had already provided sufficient information regarding these structures, further investigation of this material was not pursued. Peak C consisted of a ten residue peptide (45–54). Automated Edman degradation of 0.6 μmol of this peptide confirmed the sequence obtained from the analysis of CNBr-II.[2] Peaks D, E, and F contained mixtures of peptides 55–71, 56–71, 72–77, 31–39, and 24–28. Peak G was a pentapeptide comprising residues 40–44. Peak H contained approximately 2% of the peptide from peak G and buffer salts. Fractions in the area marked 1 in Figure 1 were pooled and analyzed. This area contained the dipeptide, Val-Lys (29–30). Lysine was obtained in 10% yield as the major amino acid in area 2 and was attributed to position 55.

Fractions comprising peaks D, E, and F were combined and the peptides were fractionated by ion exchange chromatography on DEAE-cellulose (Figure 2). Peak DE-1 consisted of a nonapeptide (31–39) and peak DE-2 was the COOH-terminal hexapeptide (Table II), both with already established sequences as shown separately.[2] Peak DE-3 was a mixture of peptides occurring in very small yield (2%) and was not studied further. For peak DE-4, which was composed of residues 55–71, COOH-terminal analysis with carboxypeptidase A and B revealed the sequence -Tyr-Phe-Val-Glu-Leu-Arg-COOH, which agreed with the data obtained by carboxypeptidase Y digestion (fraction S-1) and the sequence of CNBr-II.[2] Peak DE-5 was the only peptide containing methionine (24–28); its sequence had been well documented. Peak DE-6 contained residues 56–71. Automated Edman degradation of this peptide (0.2 μmol) conclusively established the sequence shown for CNBr-II.[2] The repetitive yield based on Pth-Ala (cycle 1) and Pth-Phe (cycle 8) was 92%. Table III provides the quantitative yields and compositions of the tryptic peptides which were isolated and analyzed.

The complete amino acid sequence of Rhesus apo A-II based upon analyses of the intact polypeptide and derived tryptic and cyanogen bromide fragments is presented in Figure 3.

Discussion

The present account describes the primary structural

analysis of the A-II protein isolated from the serum HDL of a single normolipemic male Rhesus monkey. This work was prompted by previous reports which indicated that interesting structural relationships exist between Rhesus and human A-II (Edelstein et al., 1973, 1974), and by the importance of assessing such relationships in greater detail both from the structural and phylogenetic standpoints. The strategy that dominated our approach in elucidating the sequence of Rhesus A-II was to obtain large fragments amenable to automated Edman degradation. This approach was made possible in part through the use of the enzyme <Glu-peptidase, which facilitated removal of the NH_2-terminal pyroglutamic acid residue. Sequential analysis of the unblocked polypeptide established the location of the single methionine at position 26. Cleavage with cyanogen bromide yielded a 26-residue NH_2-terminal fragment (CNBr-I) and the COOH-terminal peptide of 51 residues (CNBr-II). Automated Edman degradation of unblocked A-II and CNBr-II allowed for the assignment of 75 of the 77 residues in the A-II polypeptide. Cleavage at the single arginine residue at position 71 also provided 2 fragments (71 and 6 residues) of importance to the structural study. Sequence analysis of peptides obtained by tryptic digestion enabled us to complete the structure and, in addition, provided corroboration for sequences derived by automated Edman degradation.

The results of the present investigation establish that Rhesus A-II is a 77-residue polypeptide which exhibits a great degree of homology with the monomeric form of human A-II. Rhesus A-II, like human A-II, has pyroglutamic acid at the NH_2 terminus. Sequence analysis has delineated the positions of six amino acid replacements, most of which are conservative except for position 3 where glutamic acid replaces lysine. These mutations appear to be of little structural significance, however, at least as assessed by the results of previous circular dichroic data (Edelstein et al., 1973) and by predictive secondary structural analyses conducted according to the method of Chou and Fasman (1974a,b). The substitutions occurring in the Rhesus apoprotein were found to have a higher helical potential than the corresponding residues in the human form (Scanu et al., 1974a). This is especially significant for the alanine replacement of serine in position 40. Nonetheless, the predicted helical regions, consisting of residues 25–30 and 33–49, appear to be identical in both the human and Rhesus proteins. The significance of such observations is not readily apparent, however, and further experimentation is required to clarify this point.

We previously reported (Edelstein et al., 1973, 1974) and now confirm that Rhesus A-II, unlike the human disulfide-linked dimer, has no cysteine and thus exists in monomeric form. This mutation appears to be related to the evolutionary process. Although dimers of A-II are observed in man and in the chimpanzee (Scanu et al., 1974b) and in the orangutan and gorilla (Scanu et al., unpublished observations), the monomeric form has been invariably documented in our studies on HDL of nonhuman primates. Thus far, the analysis has included representative species of the Old and New World monkeys (Scanu et al., 1974a) as well as the gibbon, a primate which, in the phylogenetic scale, occupies a position intermediate between the Old World monkeys and the Pan family. The possible structural advantages of the A-II dimer over the monomer in terms of HDL structure and function are not obvious at present and could become a subject for future inquiry. The lipid binding capacities of free and covalently linked monomers have been shown to be similar (Scanu et al., 1974a). However, the techniques employed may not have been of sufficient sensitivity to reveal possible differences. The present studies indicate the importance of phylogenetic investigations in the study of the structure of serum lipoproteins, an approach which has already produced interesting results with other proteins.

Supplementary Material Available

Gas chromatographic patterns and amino acid analyses (six pages). Ordering information is given on any current masthead page.

References

Beckman Instruments, Beckman Bulletin No. A-TB-044.

Bennett, C. J. (1967), *Methods Enzymol. 11,* 330.

Brewer, N. B., Jr., Lux, S. E., Ronan, R., and John, K. M. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 1304.

Chou, P. Y., and Fasman, G. D. (1974a), *Biochemistry 13,* 211.

Chou, P. Y., and Fasman, G. D. (1974b), *Biochemistry 13,* 222.

Davis, B. J. (1964), *Ann. N.Y. Acad. Sci. 121,* 404.

Edelstein, C., Lim, C. T., and Scanu, A. M. (1973), *J. Biol. Chem. 248,* 7653.

Edelstein, C., Noyes, C., and Scanu, A. M. (1974), *FEBS Lett. 38,* 166.

Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1,* 80.

Fields, R. (1972), *Methods Enzymol. 25,* 464.

Hayashi, R., Moore, S., and Stein, W. H. (1973), *J. Biol. Chem. 248,* 2296.

Inagami, T., and Murakami, K. (1972), *Anal. Biochem. 47,* 501.

Jeppson, J. O., and Sjöquist, J. (1967), *Anal. Biochem. 18,* 264.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951), *J. Biol. Chem. 193,* 265.

Moore, S., and Stein, W. H. (1963), *Methods Enzymol. 6,* 819.

Noble, R. P. (1968), *J. Lipid Res. 9,* 693.

Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 244,* 5597.

Scanu, A. M., and Edelstein, C. (1971), *Anal. Biochem. 44,* 576.

Scanu, A. M., Edelstein, C., and Keim, P. (1974a), in The Plasma Proteins, 2nd ed, Putnam, F., Ed., New York, N.Y., Academic Press, p 317.

Scanu, A. M., Edelstein, C., Vitello, L., Jones, R., and Wissler, R. (1973), *J. Biol. Chem. 248,* 7648.

Scanu, A. M., Edelstein, C., and Wolf, R. H. (1974b), *Biochim. Biophys. Acta 351,* 341.

Scanu, A. M., Lim, C. T., and Edelstein, C. (1972), *J. Biol. Chem. 247,* 5850.

Scanu, A. M., Toth, J., Edelstein, C., Koga, S., and Stiller, E. (1969), *Biochemistry 8,* 3309.

Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G., and Ballantyne, D. L. (1971), *Biochemistry 10,* 4912.

Sterner, R., Noyes, C., and Heinrikson, R. L. (1974), *Biochemistry 13,* 91.

Udenfriend, S., Stein, S., Bohlen, P., and Daisman, W. (1972), in The Chemistry and Biology of Peptides,

Meienhofer, J., Ed., Ann Arbor, Mich., Ann Arbor Science Publishers, p 655.
Weber, K., and Osborn, H. (1969), *J. Biol. Chem. 244,* 4406.
Yamada, S., and Itano, H. A. (1966), *Biochim. Biophys. Acta 130,* 538.

Transfer of Xylose to Steroids by Rabbit Liver Microsomes†

Donald S. Layne,* Rosalind S. Labow, Alenka Paquet, and Denis G. Williamson

ABSTRACT: Rabbit liver microsomal preparations can transfer xylose from UDP-xylose to estrone, 17α-estradiol, and 17β-estradiol, and, in poorer yield, to diethylstilbestrol and *p*-nitrophenol. No transfer of xylose to estriol, testosterone, epitestosterone or 17α-estradiol 3-glucuronide could be demonstrated. The xyloside of [6,7-^{3}H]estrone which was formed by liver microsomes crystallized to constant specific activity with estrone β-D-xylopyranoside, the chemical preparation of which is described.

The transfer by liver microsomal preparations of nonacidic hexoses from their respective uridine nucleotides to specific hydroxyls on the estrogen molecule has been demonstrated in several species, including the human (Collins et al., 1968, 1970; Jirku and Levitz, 1972; Williamson et al., 1972; Labow et al., 1974; Labow and Layne, 1974). Rabbit liver microsomes can effect the transfer of *N*-acetylglucosamine or glucose to the 17α-hydroxyl of 17α-estradiol, provided that the 3-hydroxyl group of the steroid has previously been conjugated with glucuronic acid (Collins et al., 1968, 1970). Williamson et al. (1971) have shown that glucose or galactose, but not *N*-acetylglucosamine, can be transferred by washed rabbit liver microsomes to the phenolic 3-hydroxyl of estrone, 17α-estradiol, and 17β-estradiol, but not of estriol. These workers obtained evidence that this transfer was effected by an enzyme which was distinct from that which synthesized the 17-glycosides of 17α-estradiol 3-glucuronide, and that the formation of the 3-glucoside and the 3-galactoside was not totally dependent on the presence of UDP-glucose or UDP-galactose.

During an investigation of this reaction we obtained evidence for the formation of a conjugate of estrone when UDP-xylose was included in the incubation mixture, and the present paper presents the identification of this compound as 17-oxoestra-1,3,5(10)-trien-3-yl β-D-xylopyranoside by comparison with synthetic material, and an investigation of the specificity and characteristics of the xylosyl transferase involved.

Experimental Procedure

Materials. UDP-xylose was purchased from Sigma Chemical Co., St. Louis, Missouri. Steroids and other reagents were obtained and purified as previously described (Collins et al., 1968, 1970; Williamson et al., 1972; Labow et al., 1974). All labeled steroids contained tritium in the (6,7) positions and were obtained from New England Nuclear Corp., Boston, Mass. The materials were checked for purity by thin-layer chromatography. The specific activities used in all assays were between 30 and 40 Ci/mol.

General Methods. The procedures for the preparation of homogenates and of microsomes, thin-layer chromatography of steroid glycosides and their aglycones, the determination of protein and the assay of radioactivity in incubates and on thin-layer plates were all carried out as previously detailed (Collins et al., 1970; Williamson et al., 1972; Labow et al., 1974).

Assay of Transferase Activity. A methanol solution of 0.05 μmol of labeled steroid was evaporated to dryness in a 15-ml conical centrifuge tube fitted with a ground-glass stopper. A solution of 0.5 μmol of UDP-xylose and a microsomal suspension equivalent to 0.5 g of rabbit liver (approximately 10 mg of protein), each in 1.0 ml of 0.15 M Tris-HCl buffer, pH 7.0, were added. The total volume of buffer was brought to 3 ml, and the tubes were incubated for 30 min at 37 °C. The incubations and extraction of the media were done as described by Labow and Layne (1972), with the following modifications necessitated by the fact that the estrogen xylosides were appreciably soluble in benzene: the incubations were stopped by shaking with 5 ml of ethyl acetate. This extract was then evaporated to dryness and chromatographed in $CHCl_3$–MeOH (9:1) on silica gel H plates after the addition of standard estrone xyloside to each sample. After spraying with 2% H_2SO_4 in ethanol and heating at 110 °C for 5 min, the estrone xyloside area was scraped and counted and the percent conversion of the tritiated steroid to conjugate was calculated. When compounds other than estrone were tested as substrates, the entire plate was scraped in 1-cm sections and the radioactivity in the xyloside area was located by comparison with a control reaction incubated without UDP-xylose. The percent conversion was calculated as described above. Values were expressed either as picomoles of xyloside formed per minute per gram of liver, or per milligram of protein as determined by the method of Lowry et al. (1951).

Synthesis of Estrone β-D-Xylopyranoside. The improved Koenigs–Knorr procedure described by Conrow and Bernstein (1971) for the preparation of estrogen hexopyranosides was used to couple estrone with xylose. Estrone (2.08 g, 0.77 mmol) was reacted with 2,3,4-tri-*O*-acetyl-α-D-xylo-

† From the Department of Biochemistry, University of Ottawa, Ottawa, Canada K1N 9A9. *Received July 8, 1975.* This work was supported by Grant MT-3287 from the Medical Research Council of Canada.