bacteriophage induced DNA polymerase which was studied by using both kinetic and product size analyses (Das & Fujimura, 1977a,b).

Acknowledgments

We thank Frederic Richards for helpful suggestions during the preparation of the manuscript, William Cosina for help with preliminary studies, Sherman Weissman for SV-40 DNA fragments, and William Melchior for many thought provoking discussions.

References

Bambara, R. A., Uyemura, D., & Choi, T. (1978) J. Biol. Chem. 253, 413-423.

Chang, L. M. S. (1975) J. Mol. Biol. 93, 219-235.

Clayman, C. H., Mosharrata, E., & Faras, A. J. (1979) J. Virol. 29, 242-249.

Darling, T. L., & Reid, T. W. (1979) Nucleic Acids Res. 6, 1189-1201.

Darling, T., Albert, J., Russell, P., Albert, D. M., & Reid, T. W. (1977) J. Chromatogr. 131, 383-390.

Das, S. K., & Fujimura, R. K. (1977a) J. Biol. Chem. 252, 8700-8707.

Das, S. K., & Fujimura, R. K. (1977b) J. Biol. Chem. 252, 8708-8712.

Das, S. K., & Fujimura, R. K. (1979) J. Biol. Chem. 254, 1227-1239.

Dube, D. K., & Loeb, L. A. (1976) Biochemistry 15, 3605-3611.

Gregerson, D. S., Russell, R., & Reid, T. W. (1979) J. Gen. Virol. 43, 327-337.

Leis, J. (1976) J. Virol. 19, 932-939.

Maniatis, T., Jeffrey, A. Y., & Kleid, D. (1975) *Proc. Natl. Acad. Sci. U.S.A.* 72, 1184-1188.

McClure, W. R., & Jovin, T. M. (1975) J. Biol. Chem. 250, 4073-4080.

Riley, M., Maling, B., & Chamberlin, M. J. (1966) *J. Mol. Biol.* 20, 359-389.

Rothenberg, E., & Baltimore, D. (1976) J. Virol. 17, 168-174. Rothenberg, E., & Baltimore, D. (1977) J. Virol. 21, 168-178. Sherman, L. A., & Gefter, M. L. (1976) J. Mol. Biol. 103, 61-76.

Uyemura, D., Bambara, R., & Lehman, I. R. (1975) J. Biol. Chem. 250, 8577-8584.

Vamvakopoulos, N. C., Vournakis, J. N., & Marcus, S. L. (1977) Nucleic Acids Res. 4, 3589-3597.

Primary Structure of Murine Major Histocompatibility Complex Alloantigens: Amino Acid Sequence of the Amino-Terminal One Hundred and Seventy-three Residues of the H-2K^b Glycoprotein[†]

Hiroshi Uehara, Bruce M. Ewenstein, John M. Martinko, Stanley G. Nathenson,* John E. Coligan, and Thomas J. Kindt

ABSTRACT: The amino-terminal 173 residues of the murine histocompatibility antigen H-2Kb have been assigned by using radiochemical methodology. The complete sequence of an 86 residue glycopeptide (CN-Ib), which is one of the five major CNBr fragments of K^b, was determined by analysis of peptides obtained from digests using thrombin and V8 staphylococcal protease. Complete sequences were obtained for the three large thrombic peptides, and these were aligned by using peptides from the V8 protease digest. Alignment of the CNBr fragments was carried out by using [35S]Met-labeled peptides from a tryptic digest of the papain-cleaved H-2Kb molecule. Positive identification was possible for all the common amino acids except Asp (Asp) which was indirectly assigned and which is designated in italics. The sequence obtained in our studies was Gly-Pro-His-Ser-Leu-Arg-Tyr-Phe-Val-Thr-Ala-Val-Ser-Arg-Pro-Gly-Leu-Gly-Glu-Pro-Arg-Tyr-Met-Glu-ValGly-Tyr-Val-Asp-Asp-Thr-Glu-Phe-Val-Arg-Phe-Asp-Ser-Asp-Ala-Glu-Asn-Pro-Arg-Tyr-Glu-Pro-Arg-Ala-Arg-Trp-Met-Glu-Glu-Gly-Pro-Glu-Tyr-Trp-Glu-Arg-Glu-Thr-Gln-Lys-Ala-Lys-Gly-Asn-Glu-Gln-Ser-Phe-Arg-Val-Asp-Leu-Arg-Thr-Leu-Leu-Gly-Tyr-Tyr-(Asn)-Gln-Ser-Lys-Gly-Gly-Ser-His-Thr-Ile-Gln-Val-Ile-Ser-Gly-Cys-Glu-Val-Gly-Ser-Asp-Gly-Arg-Leu-Leu-Arg-Gly-Tyr-Gln-Gln-Tyr-Ala-Tyr-Asp-Gly-Cys-Asp-Tyr-Ile-Ala-Leu-Asn-Glu-Asp-Leu-Lys-Thr-Trp-Thr-Ala-Ala-Asp-Met-Ala-Ala-Leu-Ile-Thr-Lys-His-Lys-Trp-Glu-Gln-Ala-Gly-Glu-Ala-Glu-Arg-Leu-Arg-Ala-Tyr-Leu-Glu-Gly-Thr-Cys-Val-Glu-Trp-Leu-Arg-Arg-Tyr-Leu-Lys. These data represent the longest reported amino acid sequence determined by utilizing radiochemical methodology and provide the first extensive information on the primary structure of murine histocompatibility antigens.

he major histocompatibility complex (MHC)¹ [see reviews by Klein (1975), Shreffler & David (1975), Vitetta & Capra

(1978), Snell et al. (1976), Goetze (1977), and Festenstein & Demant (1979)] is a linked set of genes, the products of which have important roles in host reaction to antigenic stimulation. The murine MHC, which encompasses the original H-2 transplantation locus defined by Gorer (1936), is presently defined as the genetic material on chromosome 17 from the K region (nearest to the centromere) to the TLa region, a distance of about 1.5 cmorgans away.

[†] From the Department of Microbiology and Immunology and the Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461 (H.U., B.M.E., J.M.M., and S.G.N.), and the Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205 (J.E.C. and T.J.K.). Received August 7, 1979. These studies were supported in part by grants from the National Institutes of Health, AI-07289 and AI-10702 (S.G.N. and H.U.) and 5T32-CA-19931 (J.M.M.). B.M.E. was a fellow of the Cancer Research Institute.

¹ Abbreviations used: MHC, major histocompatibility complex; PTH, phenylthiohydantoin.

The classical K and D transplantation H-2 antigens (the products of K and D MHC regions) are integral, cell membrane glycoproteins of approximately 340 amino acids in length. These antigens have been shown to play important roles in the stimulation and specificity of T lymphocyte responses to virally induced and other cell surface antigens (Shearer et al., 1975; Doherty et al., 1976; Paul & Benacerraf, 1977). It is clear that an understanding of the function of the H-2 molecule in molecular terms can be approached only when its intramolecular organization has been elucidated relative to its known properties. Thus, it is important to define those regions involved in binding to the cell membrane, intracellular interactions (possibly with cytoskeletal elements), and interactions with the β_2 -microglobulin subunit and, most important, those regions involved in the alloantigenic activity and in the T-cell mediated associative recognition function.

Shimada & Nathenson (1969) established that the K and D gene products of the mouse H-2 MHC could be solubilized from the cell surface by digestion with the proteolytic enzyme papain and were glycoproteins with molecular weights of approximately 37 000. Later studies on the intact H-2 glycoproteins solubilized by the detergent NP-40 (Schwartz et al., 1973) indicated that the soluble papain-cleaved fragments comprised the amino-terminal 80% of the native molecules (Ewenstein et al., 1976). The K/D molecules occur on the cell surface in a 1:1 association with a small nonpolymorphic protein, β_2 -microglobulin, which has a molecular weight of 11 500 (Rask et al., 1974; Silver & Hood, 1974; Natori et al., 1975).

Because of the difficulty in obtaining significant amounts of the H-2 molecules in purified form, most structural studies have utilized intrinsically radiolabeled material obtained from cells cultured in the presence of radioactive amino acids. When selected K and D products of different H-2 haplotypes were examined by peptide mapping studies, an extensive variability (sharing of only 30-40% of the peptides) was noted to exist between K and D products from mice having different H-2 haplotypes (Brown et al., 1974). Such findings correlated with the previously noted H-2 antigenic variability [reviewed in Klein (1975) and Snell et al. (1976)].

Primary structural studies on the MHC product encoded by the H-2K gene of the H-2b haplotype have been initiated also by using radiochemical methods. Five major CNBr fragments of H-2Kb were isolated, and their N-terminal sequences were determined. They were provisionally aligned as follows: CN-IIIn (23 residues) \rightarrow CN-IIIa (29 residues) \rightarrow CN-Ib (86 residues) \rightarrow CN-Ia (90 residues) \rightarrow CN-Ic (56 residues) (Ewenstein et al., 1978; Coligan et al., 1978, 1979). This report documents the assignment of the NH₂-terminal 173 residues of the H-2Kb (H-2.33) glycoprotein and further demonstrates that complete amino acid sequence analysis of radiolabeled molecules can be accomplished by strategies similar to those applied to unlabeled proteins.

Materials and Methods

Antisera, Radiolabeling, and Preparation of Cell Extracts. The primary alloantiserum used to isolate the H-2Kb glycoprotein was prepared in (B10.D2 × A)F₁ mice by injection with spleen cells from HT1 [H-2d(Kd,Dd) × H-2d(Kb,Dd) anti H-2d(Kb,Dd)] as previously described (Ewenstein et al., 1976). The incorporation of radiolabeled amino acids was carried out by a minor modification (R. Nairn and S. G. Nathenson, unpublished data) of previously published techniques (Brown et al., 1974). In brief, EL-4-BrdU cells, harvested in rapid growth phase, were resuspended at $\sim 5 \times 10^6/\text{mL}$ in Dubecco's modified Eagles medium (Grand Island Biological Co.,

Grand Island, NY) containing 10% fetal calf serum and one or more radiolabeled amino acids. Cells were cultured for 8 h at 37 °C in 5% CO₂ and subsequently isolated and extracted with 0.5% Non-Idet P-40. Radioactive amino acids were purchased from New England Nuclear Corp., Boston, MA, except for [³H]Asn which was purchased from Amersham/ Searle Corp., Arlington Heights, IL.

Isolation of the H-2K^b Glycoprotein. The H-2K^b glycoprotein was isolated by indirect immune precipitation of lentil-lectin purified NP-40 extracts of cells, followed by Sephadex G-75 column chromatography as described previously (Ewenstein et al., 1978). The β_2 -microglobulin was isolated as a byproduct during the procedure, and the NH₂-terminal sequence for approximately 30 steps was determined in order to establish relative levels of radioactivity for the incorporated amino acids. These data also provided information on the extent of biosynthetic interconversion of radiolabeled amino acids [cf. footnotes of Table I and Coligan et al. (1978)].

Isolation of CNBr Fragments. The CNBr fragments of the H-2K^b molecule were obtained as described by Ewenstein et al. (1978) except for fragments CN-IIIn and CN-IIIa. These fragments, previously separated by ion-exchange chromatography, were isolated by using a 1.5 × 220 cm Sephadex G-50 (superfine) column developed in 1 M HCOOH. The use of this technique permitted the composite fragments of pool III (i.e., CNBr fragments IIIn, IIIa, and IIIc) to be purified in one step.

Automated Amino Acid Sequence Analysis. Automated Edman degradation was performed with a Beckman sequencer (Model 890-C) using Beckman Peptide Program No. 102974 in the presence of several milligrams of carrier protein, bovine serum albumin, and/or horse heart cytochrome c (Sigma Chemical Co., St. Louis, MO). Sequence analysis of peptides smaller than 50 residues was performed in the presence of Polybrene (Aldrich Chemical Co., Milwaukee, WI) as described by Klapper et al. (1978).

In many cases in which peptides were radiolabeled with only one amino acid, the total butyl chloride extract was counted. Radioactivity significantly above background indicated the presence of that amino acid at that step. In these cases, the assignments were considered final only when the same results were obtained from sequence analysis of either a duplicate sample of the peptide or an overlapping peptide obtained by a different proteolytic cleavage method. For peptides containing more than one radiolabeled amino acid as well as for many peptides labeled by single amino acids, an aliquot of the butyl chloride extract was counted and the remainder of those samples having radioactivity were evaporated. The thiazolinone amino acids in these extracts were converted to phenylthiohydantoin (PTH) derivatives by heating to 80 °C in 1 N HCl for 10 min followed by ethyl acetate extraction of the amino acid PTH derivative. An exception to this procedure was for glycine which was converted as previously described (Coligan et al., 1979). The presence of a significant amount of radioactivity in the aqueous phase after ethyl acetate extraction was taken to indicate the presence of Arg-PTH or His-PTH at that step. Radiolabeled amino acid PTH derivatives were identified by the presence of radioactivity in the position of nonradioactive amino acid PTH derivative standards eluting during high-pressure liquid chromatography (LC) and/or thin-layer chromatography (TLC). Details of these procedures were described previously (Coligan et al., 1979; Gates et al., 1979).

Thrombin Digestion. Thrombin was obtained by partial purification of thrombin (B grade, Lot No. 703053, Calbio-

308 BIOCHEMISTRY UEHARA ET AL.

chem, La Jolla, CA) as described by Lundblad (1971). Active fractions eluted from SP-Sephadex C-50 columns were dialyzed against 0.01 M Tris-HCl, pH 7.5 (containing 0.05 M NaCl), and stored at –80 °C without further purification. The esterase activity of the preparation against p-tosyl-1-arginine methyl ester (TAME) as determined by the spectrophotometric method of Hummel (1959) showed a ΔA_{247} /(min·mL) equal to 1.6 at 23 °C. CNBr fragment Ib (CN-Ib) labeled with radioactive amino acids (mixed with 0.5–5 mg of carrier cytochrome c and/or bovine serum albumin) was dissolved in 0.7–0.9 mL of 0.1 M NH₄HCO₃ and incubated at 37 °C for 20 h with 0.09–0.15 mL of thrombin solution. The reaction was halted by the addition of 0.10 mL of 88% HCOOH and applied directly on a column of Sephadex G-50 (Superfine) which was equilibrated to 1 M HCOOH.

Trypsin Digestion. Trypsin digestion and ion-exchange chromatography were carried out as described previously (Brown et al., 1974) using TPCK-trypsin (Worthington Biochemicals, Freehold, NJ).

Digestion by Staphylococcus aureus V8 Protease. S. aureus protease V8 (Miles Laboratories, Inc., Elkhart, IL) was used without further treatment. CN-Ib (containing 2 mg of carrier cytochrome c) was dissolved in 0.9 mL of 1 M NH₄HCO₃ and 2 mM EDTA, and 0.1 mL of protease V8 solution (1 mg/mL in water) was added. The reaction was halted by the addition of 0.1 mL of 88% HCOOH, and the sample was applied directly to a Sephadex G-50 column, equilibrated in 1 M HCOOH.

Results

Incorporation of Radioactivity into H-2Kb. The H-2Kb glycoprotein was biosynthetically labeled by short-term culture of EL-4 cells in the presence of either individual or selected mixtures of ³H- or ³⁵S-labeled amino acids and then purified by LcH-Sepharose affinity chromatography and immune precipitation (Table I). The percentage of radioactivity in the culture incorporated into washed cells or immune precipitates of H-2Kb for each of the radiolabeled amino acids varied in a reproducible way, and chemically similar acids were incorporated with similar efficiencies. For example, amino acids with basic or aromatic long-chain aliphatic side groups were efficiently incorporated by the cells whereas lower percentages of acidic amino acids and their amides were incorporated (Table I). At present, all of the common amino acids except Asp have been incorporated in amounts sufficient to permit structural studies.

In the majority of cases, radioactivity in the isolated H-2K^b glycoprotein was associated solely with the amino acid residue which was added in radiolabeled form to the culture medium. However, a significant degree of biosynthetic interconversion was observed with a few of the amino acids, as noted in the footnotes to Table I. In these instances the proportion of radioactivity incorporated into each of the amino acids was initially determined by NH₂-terminal sequence analysis of the β_2 -microglobulin obtained as a byproduct in the purification of H-2K^b. The relative activities found were in all cases consistent with data for pure H-2K^b peptides. This interconversion is discussed elsewhere in greater detail (Coligan et al., 1979).

When groups of radioactive amino acids were utilized, it was found that the variation in incorporation efficiencies among the amino acids hampered the identification of the poorly labeled amino acid residues. This problem was partially circumvented by adjusting the relative specific activities through the supplementation of the multiply labeled H-2Kb molecule, or its subfragments, with singly labeled preparations.

Table I: Comparative Incorporation Efficiency of Different Radioactive Amino Acids^a

amino acids ^b	% incorpn into EL-4 cells ^c	% incorpn into H-2K ^b immune pre- cipitates ^{c,d}
[³ H] Arg	78	0.12
[³H] Tyr	63	0.07
[³H] Leu	53	0.03
[³⁵ S] Met	50	0.012
[³H] Lys	48	0.015
[35S] Cys	36	0.062
[³ H] Met	31	0.012
[³H] Ser ^e	29	0.011
[3H] Val	25	0.015
[³H] Gly ^f	22	0.009
[³ H] Pro	18	0.008
[³H]Gln ^g	17	0.01
[3 H] A sn	15	0.004
[³H] Ala	13	0.003
[³H]Glu ^h	11	0.005
[³H] Asp	2.5	
[³H]FLKY	42	0.028
[³H] HITVW	28	0.042

^a Experiments usually involved addition of 5-10 mCi of a given amino acid to about 3 × 108 cells in 60 mL of labeling medium (see Materials and Methods). b Specific activities of radioactive amino acids used are as follows (Ci/mmol): L-[2,3-3H]arginine, 15-25; L-[ring-2,6-3 H] tyrosine, 40-50; L-[4,5-3 H(N)] leucine, 50-60; L-[35 S] methionine, 600-800; L-[4 ,5- 3 H(N)] lysine, 72; $L-[{}^{35}S]$ cystine, 300-600; $L-[{}^{3}H]$ methionine, 6.5; $L-[{}^{3}H(G)]$ serine, 2.8; L-[2,3,43H] valine, 5-11; [2-3H] glycine, 15; L-[2,3-3H] proline, 27; L-[3,4-3H(N)] glutamine, 43; L-[3H]asparagine, 0.23; L-[3-3H] alanine, 15-50; L-[2,3-3H] glutamic acid, 51; L-[2,3-3H(N)] aspartic acid, 15-20; L-[ring-2,6-3H(N)]phenylalanine, 55; L-[3-3H]histidine, 10-20; L-[4,5-3H(N)]isoleucine, 95-115; L-[3H(G)] threonine, 2.5; L-[side chain-2,3-3H-(N)] tryptophan, 15-20. CPercentage of radioactivity added to culture. Data are the average from several preparations. d Immune precipitates contained β_2 -microglobulin in addition to H-When [3H] Ser is incorporated, 10% of radioactivity is in Gly. f When $[^3H]$ Gly is incorporated, 35% of radioactivity is in Ser. [¶] When [³H] Gln is incorporated, 17% of radioactivity is in Glu and 7% is in Pro. ^h When [³H] Glu is incorporated, 40% of radioactivity is in Pro.

Sequence Studies of CN-Ib. One of the peptides obtained from CNBr digestion of the Kb molecule was CN-Ib, a glycopeptide spanning positions 53-138. Thrombin, a proteolytic enzyme of the vertebrate blood clotting system (Seegers, 1967), was used to generate three major fragments (θ 1, θ 2, and θ 3) of CN-Ib which were particularly useful for automated sequence analysis. The fragments were isolated by gel filtration on a column of Sephadex G-50 (superfine) (Figure 1) and the NH₂-terminal sequence of each peptide was determined. Comparison of the amino acid sequence of $\theta 2$ to that of the intact fragment revealed that this peptide represented the NH₂-terminal 27 residues of CN-Ib. Similar analysis revealed that θ 1 begins at position 28. Partial amino acid sequence studies of peptide $\theta 3$ indicated that this peptide was not derived from the NH₂-terminal 40 residues of CN-Ib. The presence of [3 H]Met in $\theta 3$ indicated that this was the COOH-terminal peptide.

It was found that the elution positions of the thrombic peptides were highly reproducible, and no additional peaks were observed in any of the different separations. Quantitation of the radioactivity recovered in each of the thrombic peptides from the eleven different amino acid labeling groups is summarized in Table II. The techniques for determination of amino acid composition of radiolabeled peptides have not yet been perfected, but the data contained in Table II provide useful insight concerning the composition of the thrombic

Table II: Recovery of Radioactivity in CN-Ib Thrombic Peptides for Different Radiolabeled Amino Acids

multi- label. group	radioact. AA's ^a	$\theta 1$			θ2			θ3		
		radioact.		AA compo-	radioact.		AA compo-	radioact.		AA compo-
		cpm ^b	% c	sition ^d	cpm ^b	% ^c	sition ^d	cpm ^b	% ^c	sition ^d
I	FLKYR	21 700	31.8		15 500	22.7		20 000	45.5	
II	HITVW	24 000	33.5		23 000	32.1		16 000	34.4	
III	EΡ	800	14.3	1:0	3 400	60.9	6:1	900	24.8	1:0
IV	GS	10 000	64.7	6:4	3 000	19.4	2:1	1600	15.9	2:0
V	SG	16 000	72.1	4:6	4 500	20.3	1:2	1 100	7.6	0:2
VI	QEP	1 000	21.8	2:1:0	2500	54.6	3:6:1	700	23.5	2:1:0
	Ā	<100	0	0	700	19.4	1	1 900	81.6	4
	P	<100	0	0	1 900	100	1	<100	0	0
	N	2600	36.9	1	2 300	32.6	1	1 400	30.5	1
	С	4 800	46.4	1	<200	0	0	3600	53.6	1
	M	<100		0	<100	0	0	800	0	1

^a Amino acids are identified by the single letter code (IUPAC-IUB Commission on Biochemical Nomenclature). Relative activity for preparations radiolabeled with multiple amino acids was reported previously (Coligan et al., 1978). ^b Radioactivity (cpm) recovered in peptide. ^c Percent = (radioactivity in each thrombic peptide)/(total recovered radioactivity) \times 100. The radioactivity in θ 3 was corrected for loss during purification. ^d The amino acid composition as determined by sequence analysis.

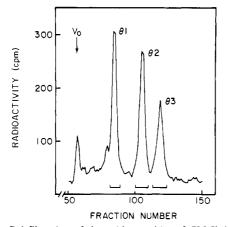


FIGURE 1: Gel filtration of thrombic peptides of CN-Ib labeled in [3 H]HITVW on a column (0.9 × 190 cm) of Sephadex G-50 (superfine) in 1.0 M HCOOH. The flow rate was 2.2 mL/h and 0.75-mL fractions were collected. The peptides were detected by sampling each fraction for radioactivity; pools were made as indicated. V_0 indicates void volume of the column.

peptides of CN-Ib. For example, only CN-Ib- θ 1 and CN-Ib- θ 3 contain half-cystine residues while Met is present only in θ 3. In addition, the radioactivity recovered in each of the thrombic peptides appears to accurately reflect the relative amino acid compositions of these peptides as determined by amino acid sequence analysis. Thus, CN-Ib- θ 1 appears to be rich in Gly and Ser, and CN-Ib- θ 2 contains large numbers of Glu and Gln residues.

Amino Acid Sequence of Thrombic Peptides. Automated sequence analysis was performed on each of the thrombic peptides derived from CN-Ib. The sequence analysis of CN-Ib- θ 2 confirmed many of the amino acid assignments in the NH₂-terminal 27 positions of fragment CN-Ib. One additional assignment, Asn-19, was made. Sequence analysis of CN-Ib- θ 1 provided sequence data continuous through residue 59 of the CNBr fragment. The amino acid sequence of residues 60-86 was obtained from the thrombic peptide θ 3. The radioactivity recovered in each step during sequence analyses of these peptides is summarized in Figure 2. The COOHterminal residues of $\theta 1$ and $\theta 2$ were found to be Arg, in agreement with the known proteolytic specificity of thrombin. Sequence analysis of peptide CN-Ib-θ3 labeled with [³H]Met yielded radioactivity at steps 26 and 27, consistent with the tentative assignment of CN-Ib- θ 3 as the COOH-terminal peptide of CN-Ib. It was not uncommon in our studies to

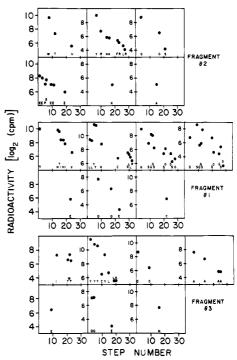


FIGURE 2: Radioactivity recovered in each step of the amino acid sequence analyses of thrombic peptides of CN-Ib radiolabeled with different amino acids.

obtain significant amounts of the COOH-terminal amino acid at the penultimate position during sequence analysis. Subsequent sequence analysis of tryptic peptides localized the Met to position 27 (see below).

Analysis of Tryptic Peptides from CN-Ib. For confirmation of some amino acid assignments, [3H]SIGHTVW- and [3H]FLKYR-labeled preparations of CN-Ib were cleaved by trypsin, and the resultant peptides were chromatographed by cation-exchange column chromatography (Figure 3). The identity of each tryptic peptide was established by sequence analysis. In two instances a single peak of radioactivity was found to contain two peptides. Thus, the amino acid sequence of the peak eluting at pH 3.96 was determined to be -Y-HY/TI--, thereby implying that this peak contained the two peptides CN-Ib-T7 and -T9 (see Table III). Similarly, the radioactive peak eluting at pH 4.24 was found to contain the peptides CN-Ib-T1 and -T4. The breakthrough peak, which was further fractionated by Sephadex G-50 gel filtration in

Table III: Identification of Amino Acids in CN-Ib trypsin NH2-terminal thrombin label group^b label assigned label group^b ID group b ID ID peptide frag. no. residuea peptide III LC LC T1(4.24) E $\theta 2$ III, VI LC LC 2 LC VI VÍ Q III, VI LC E III IV, V S, III LC LC LC 4 V LC G III 5 P E ΙΙΙ III 6 7 8 LC, TLC LC, TLC LC LC LC LC I II I Y S, I II W Π 9 Ε III LC III LC I LC 10 R S I T2(4.03) E III LC III 11 LC, TLC LC T Π LC II LC 12 II Q K VI 13 LC LC, TLC I LC Ι 14 S, I LC LC LC A K G S 15 S T3(4.53) I LC 16 S, I LC, TLC T T4(4.24) ÍV, V V 17 N E 18 S Ш III 19 LC 20 21 22 23 24 25 26 27 28 29 30 Q VIS F v IV, V LC LC LC LC I LC, TLC I Ι S2 R I Ι V II T5(4.32) II LC LC LC LC I L S I LC LC LC LC R **S2** Ι I LC LC LC II LC T6 $\theta 1$ T II Π Ι L S I S I L I IV, V 31 G LC LC LC 32 33 34 Y Y S I S LC I VI LC 35 Q S K VI 36 37 IV2, V I LC S LC Ι IV, V IV, V IV2, V T7_(3.95) 38 G S H 39 40 LC LC LC II 41 II LC LC LC T II Π 42 II 43 I Η 44 Q V VI VI LC LC II 45 46 I Η IV, V IV, V 47 S G C E V 48 LC **S**2 49 S 50 III, VI LC 51 52 53 54 55 56 57 58 59 60 61 II IV, V G LC S IV, V IV, V G LC LC I R L L R G Y Q Q Y LC LC LC T8(4.87) I I I LC I T9(3.97) ĪV, V θ 3 LC LC LC 62 63 64 65 66 67 68 VI LC LC LC LC VII LC Ι T10(3.80) A Y **S**2 I LC I G C IV, V S 69 70 71 72 73 74 S LC LC LC Y

LC LC

LC

LC

Ι

II

S2

A L

frag. no.	assigned residue ^a	NH ₂ -terminal		thrombin			try psin		
		label group ^b	ID	peptide	label group ^b	ID	peptide	label group ^b	ID
7.5	N				S	LC			
76	E				III, IV				
77									
78	L				I	LC			
79	K				I	LC			
80	T				II	LC	T11 _(3.56)	II	LC
81	\mathbf{w}				II			II	LC
82	T				II	LC		II	LC
83	Α				S2	LC			
84	Α				S2	LC			
85									
86	M				S				

^a The standard one-letter symbols (20) for amino acids are used. ^b The labeling group indicates whether a given radiolabeled amino acid was incorporated singly, as denoted by S, or as one of several radiolabeled amino acids, as denoted by I-VI (see Table II). Numbers following S indicate the number of replicate single-label runs. All radiolabeled amino acid PTH derivatives in multilabeled preparations and many in radiolabeled preparations were identified by high-pressure liquid chromatography (LC) and/or thin-layer chromatography (TLC) as indicated.

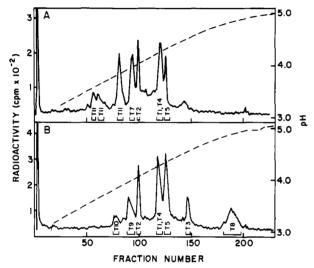


FIGURE 3: Ion-exchange chromatography of tryptic peptides of CN-Ib radiolabeled in [3H]SIGHTVW (A) and [3H]FLKYR (B). Details of chromatographic conditions were described by Brown et al. (1974). The peptides were pooled as indicated. The peptides not bound to the column (fractions 2–5) were further purified by gel filtration on Sephadex G-50 (superfine) in 1 M HCOOH. The sequence analysis for each peptide is shown in Figure 5 and documented in Table III.

1 M HCOOH, was found to contain the carbohydrate-containing peptide as determined by labeling studies utilizing [³H]fucose. The sequence data obtained from tryptic peptides of CN-Ib confirmed assignments made by sequence analysis of both CN-Ib and its thrombic peptides. These data are summarized in Table III.

Alignment of CN-Ib Peptides. For establishment of the unambiguous alignment of the thrombic peptides, CN-Ib radiolabeled with Arg, Leu, and Tyr was digested with staphylococcal protease V8. The reaction mixture was applied to a Sephadex G-50 column, resulting in the fractionation shown in Figure 4. Two peptides (Ib-V8-1 and Ib-V8-2) were isolated and their amino acid sequence was determined. The amino acid sequence of Ib-V8-1 was determined to be LR-LL-YY--, thus confirming the alignment of CN-Ib-θ2 to CN-Ib-θ1, which had been established from the NH₂-terminal sequence analyses. The sequence of a Ib-V8-2 was found to be ----RLLR-Y--Y-Y--, providing the evidence for alignment of CN-Ib-θ1 to CN-Ib-θ3 (see Figure 5).

Summary of Amino Acid Sequence Data for CN-Ib. The complete amino acid sequence of CN-Ib derived from throm-

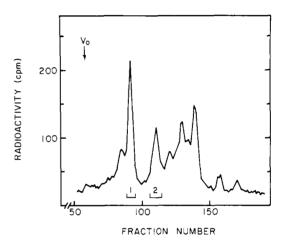


FIGURE 4: Gel filtration of staphylococcal protease V8 peptides of CN-Ib labeled in [³H]RLY on a column of Sephadex G-50 (superfine) as described in Figure 1.

bic, tryptic, and V8 protease peptides is shown in Figure 5 and documented in Table III. Positions 25, 54, 67, 70, 77, and 85 are tentatively assigned as Asp (D) because, if another amino acid occurred at these positions, it would most likely have been detected. (Asp is the amino acid that has not yet been incorporated successfully into the H-2K^b molecule.)

It had been previously suggested that the carbohydrate units of H-2K^b were attached via Asn linkages (Nathenson & Muramatsu, 1971). Tryptic glycopeptide T6 (residues 28–37) labels with Asn; however, no Asn was observed in the sequence analysis of this peptide. Since residue 34 is the only undetermined position in this peptide, Asn has been assigned to this position, and it is presumed that this is the linkage site of the carbohydrate moiety. Further suggestive evidence that this is the site of carbohydrate attachment comes from the fact that the sequence around residue 34 would then be Asn-Gln-Ser, which is a recognition sequence for glycosylation (Marshall, 1972).

CNBr Fragments IIIn and IIIa. In a previous report the assignments of all amino acids except for Asn and Asp were made in CNBr fragments IIIn and IIIa. No unassigned positions remained in CNBr IIIn. Since that report, K^b has been labeled with [³H]Asn and the CNBr fragments have been isolated. Consistent with previous data, CN-IIIn contained no Asn. Sequence determination of CN-IIIa showed Asn at position 19.

312 BIOCHEMISTRY UEHARA ET AL.

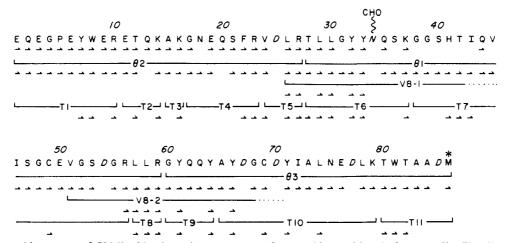


FIGURE 5: Amino acid sequence of CN-Ib. Numbers above sequence refer to residue positions in fragment Ib. First line of arrows depicts residues identified by NH_2 -terminal sequence analysis of CN-Ib. Other arrows depict the sequences for thrombic peptides (θ), staphylococcal protease V8 peptides (V8), and tryptic peptides (T). The asterisk indicates the residue identified from the sequence of a tryptic peptide derived from the papain-cleaved molecule.

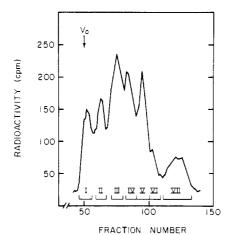


FIGURE 6: Gel filtration of tryptic peptides from the papain fragment of H-2Kb labeled in [3H]HITVW/[35S]M on a column (0.9 × 110 cm) of Sephadex G-25 (superfine) in 1 M HCOOH. Flow rate was 2.1 mL/h and 0.65-mL fractions were collected. The amount of radioactivity contained in 0.03-mL samples of each fraction was determined and pools were made as indicated. V_0 indicates void volume of the column.

Positions 6, 7, 14, and 16 of IIIa remain unassigned. On this basis, presumptive assignments of Asp were made for these residues.

Alignment of CNBr Fragments of K^b. The alignment of CNBr fragments for the NH₂-terminal portion of the H-2K^b molecule was accomplished by the identification of Met-containing peptides obtained from tryptic digestion of the papain-treated H-2K^b glycoprotein. For this determination, molecules labeled in [³H]His, -Ile, -Thr, -Val, -Trp, and -Tyr and [³⁵S]Met were used. After reduction and alkylation, the papain-cleaved molecule was digested with trypsin, and the tryptic peptides were fractionated by Sephadex G-25 column chromatography (Figure 6). Pools III and VII obtained by this procedure were further purified by cation-exchange chromatography and two Met-containing tryptic peptides (III-1 and VII-1, respectively) were isolated (Figure 7).

The partial amino acid sequence of tryptic peptide III-1 was found to be WM-----YW-, consistent with the alignment of CN-IIIa to CN-Ib. The amino acid sequence of peptide VII-1 was determined to be TWT---M--IT-, thus establishing the alignment of CN-Ib to CN-Ia. These two overlap peptides are indicated in Figure 8 by a dotted line. The alignment of CN-IIIn to CN-IIIa initially determined by the

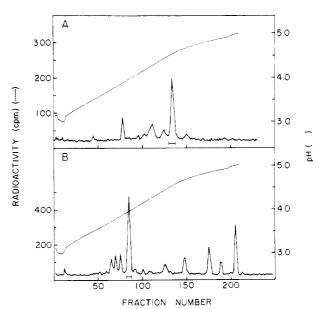


FIGURE 7: Ion-exchange chromatography of tryptic peptides of pool VII (panel A) and pool III (B) of Figure 6. Chromatographic conditions have been described by Brown et al. (1974). Fractions containing ³⁵S radioactivity were pooled as indicated by bars, yielding pure peptides VII-1 (panel A) and III-1 (panel B).

NH₂-terminal sequence of the intact H-2K^b molecule was reconfirmed by the isolation of a Met-containing peptide during this procedure with the sequence YM-V-YV- (see Figure 8).

Amino Acid Sequence of Residues 1-173 of H-2K^b. The amino acid sequence of NH₂-terminal 173 residues of K^b was thus established, and data supporting this sequence are shown in Figure 8.

Discussion

The amino acid sequence has been determined for residues 1–173 of the radiolabeled H-2K^b molecule by automated sequence analysis of CNBr fragments and of peptides proteolytically prepared from these fragments by tryptic, thrombic, and V8 protease digestion. Whereas radioactive sequence analysis has been used previously for partial structure studies on small proteins such as hormones (Jacobs et al., 1974) and viral peptides (McKean et al., 1974), this sequence is the longest reported continuous sequence for a protein determined by radiochemical methodology and provides the first extensive

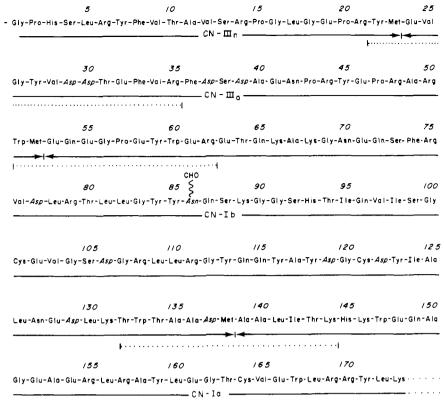


FIGURE 8: Amino acid sequences of residues 1-173 of the H-2K^b molecule. The dotted lines below the sequence show the peptides used to align the CNBr fragments.

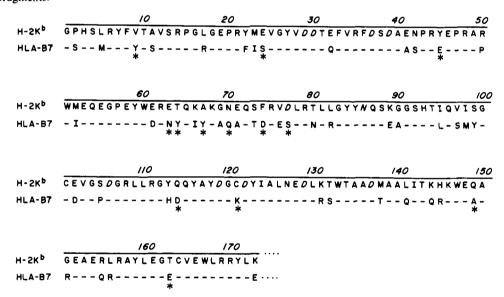


FIGURE 9: Comparison of the amino acid sequences of H-2K^b and HLA-B7 (Strominger et al., 1979). Dashed lines indicate identity; positions of different residues identified are as indicated. The asterisk indicates amino acid interchanges which require more than a single base change.

information on a murine histocompatibility antigen.

The NH₂-terminal portion of the molecule is important because it represents the majority of the papain-derived molecule (approximately positions 1–270), which has been shown to carry the H-2K^b alloantigenic determinants but lacks the membrane binding region. Interest in this region also stems from preliminary data (Nathenson et al., 1979) suggesting that the K^b glycoproteins from the MHC mutant K^b mouse lines (Bailey & Kohn, 1965; Egorov, 1967; Melvold & Kohn, 1976) have discrete amino acid changes in the NH₂-terminal half of the molecule, specifically at amino acid positions 155 (bm1) and 77 and 89 (bm3). These small discrete differences correlate with remarkable alterations in biological reactivity between parent and mutant strains.

Although many of the techniques used in radiosequence analysis are similar to those used for conventional studies, there are differences, especially in criteria used for assignment of amino acid residues. In this study, the following guidelines were generally followed.

- (a) For analysis of peptides labeled with a single amino acid, assignments were made at positions giving significant radio-activity in the butyl chloride extract if verified by LC analysis or, in cases where the total butyl chloride extract was counted, by similar results in duplicate analyses of the peptide.
- (b) If the same amino acid appeared at more than one step in a sequencer run, an identification was considered valid only if the yield at these steps fitted the expected yield as determined by the characteristic logarithmic repetitive yield curve.

314 BIOCHEMISTRY UEHARA ET AL.

Recoveries of radioactivity for other residues were also required to be related to this curve as determined by the ratios of specific activities of the amino acids involved as previously determined by analysis of β_2 -microglobulin. Repetitive yields ranged from 89 to 93% and were reasonably constant in multiple analyses of any given peptide.

(c) When possible, assignments were made only after a given amino acid was observed at a certain position in more than one analysis (see Table III). This requirement was fulfilled by the isolation and sequence analysis of peptides from different digests of the same region of the molecule or by duplicate analyses of the same peptide.

Several features of the amino acid sequence of residues 1-173 warrant special mention (see Figure 5). The aminoterminal 50 residues are relatively rich in Arg while no Lys residues have been assigned for this region. Short sequences containing a high percentage of Glu residues are found in residues 53-63. Three half-cystine residues are present in positions 101, 121, and 164. Preliminary studies from our laboratory (J. Martinko, unpublished observation) suggest that a disulfide bridge from Cys-101 to Cys-164 is present, and such data agree with the information on HLA molecules which shows homologous Cys at residues 101 and 164. The attachment site of the carbohydrate chain was assigned to position 86. Only one carbohydrate chain is found within the NH₂-terminal 173 residues; however, preliminary evidence suggests that the second carbohydrate is at position 176 (J. Martinko, unpublished observation).

Extensive sequence comparisons of K^b with other murine MHC molecules are not yet possible due to lack of comparable sequence information. However, NH₂-terminal regions of several H-2 molecules, e.g., H-2K^k, H-2K^q, and H-2D^d, have been obtained from our and a number of other laboratories (Vitetta & Capra, 1978), and approximately 85% homology is observed for the positions that can be compared. We are presently analyzing the D^b, K^d, and D^d molecules by the radiochemical approach. Thus far, the D^d sequence is the most advanced and for about 110 residues shows a sequence homology with K^b of about 85% (R. Nairn, unpublished observation).

Studies on the primary structure of HLA molecules (Strominger et al., 1979) allow an extensive comparison of the amino acid sequences of HLA-B7 to that of H-2Kb. Comparison of the amino acid sequence of residues 1-173 shows that 122 of the 173 positions (71% homology) are identical (Figure 9). The majority of amino acid differences between the two molecules can be attributed to single nucleotide base substitution in the genome, and thus far only 65 base substitutions are required to interconvert the HLA-B7 and the H-2K^b amino acid sequences in the 173 positions compared. Of particular interest with regard to potential functional regions is the finding that the interchanges between HLA-B7 and Kb do not distribute equally within the regions compared; it is possible to point out the presence of a major cluster of amino acid sequence differences between the two molecules. Thus, in region 61-82, of 22 positions only 8 positions are identical (36% homology). Of importance also is that a number of interchanges between residues 61-82 require two or three base substitutions. It is intriguing that the differences noted above between the bm3 mutant and the parent Kb molecules (positions 77 and 89) (Nathenson et al., 1979) occur in or close to this "diversity" region.

Of the 10 Asp assignments in the H-2K^b sequence, which are presumptively made as described under Results, all but Asp-77 are identical with the Asp assignments in HLA-B7.

However, Asp is present at 77 in the HLA-A2 molecule. Additional similarities between the murine and human MHC molecules in the first 173 residues are the presence in both of a carbohydrate chain at position 86 and the identical location of two of the three cysteine residues. The similarities in many of the chemical features as well as the sequence of the two molecules support the concept that significant similarity will be found for the secondary and tertiary structures of the histocompatibility antigens from various species. However, a more precise evaluation of this and other questions relating to the biological activity of histocompatibility antigens can be approached when more extensive structural studies of these molecules are completed.

Acknowledgments

We thank M. Flocco, D. McGovern, S. Schumacher, Pat Spinella, and R. Valas for expert technical assistance and C. Whelan for expert secretarial assistance. We thank Drs. R. Nairn and K. Yamaga for their stimulating discussions.

References

Bailey, D. W., & Kohn, H. I. (1965) Genet. Res. 6, 330.Brown, J. L., Kato, K., Silver, J., & Nathenson, S. G. (1974) Biochemistry 13, 3174.

Coligan, J. E., Kindt, T. J., Ewenstein, B. M., Uehara, H., Nisizawa, T., & Nathenson, S. G. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 3390.

Coligan, J. E., Kindt, T. J., Ewenstein, B. M., Uehara, H., Martinko, J. M., & Nathenson, S. G. (1979) Mol. Immunol. 16, 3.

Doherty, P. D., Blanden, R. Y., & Zinkernagel, R. M. (1976) Transplant. Rev. 29, 89.

Egorov, I. (1974) Immunogenetics (N.Y.) 1, 97.

Ewenstein, B. M., Freed, J. H., Mole, L. E., & Nathenson,S. G. (1976) *Proc. Natl. Acad. Sci. U.S.A.* 73, 915.

Ewenstein, B. M., Nisizawa, T., Uehara, H., Nathenson, S. G., Coligan, J. E., & Kindt, T. J. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 2909.

Festenstein, H., & Demant, P. (1979) HLA and H-2 Basic Immunogenetics, Biology and Clinical Relevance, Edward Arnold, Great Britain.

Gates, F. T., Coligan, J. E., & Kindt, T. J. (1979) Biochemistry 18, 2267.

Goetze, D., Ed. (1977) in *The Major Histocompatibility* System in Man and Animal, Springer-Verlag, Berlin and New York.

Gorer, P. A. (1936) Br. J. Exp. Pathol. 17, 42.

Hummel, B. C. W. (1959) Can. J. Biochem. Physiol. 37, 1393.
Jacobs, J. W., Kempner, B., Niall, H. D., Habener, J. F., & Pofts, J. J., Jr. (1974) Nature (London) 249, 155.

Klapper, D., Wilde, C. W., & Capra, J. D. (1978) Anal. Biochem. 85, 126.

Klein, J. (1975) Biology of the Mouse H-2 Complex, Springer-Verlag, New York.

Lundblad, R. L. (1971) Biochemistry 10, 2501.

Marshall, R. D. (1972) Annu. Rev. Biochem. 41, 673.

McKean, D. J., Peters, E. H., Waldby, J. I., & Smithies, O. (1974) *Biochemistry 13*, 3048.

Melvold, R. W., & Kohn, H. I. (1976) *Immunogenetics (N.Y.)* 3, 185.

Nathenson, S. G., & Muramatsu, T. (1971) in *Glycoproteins* of *Blood Cells* and *Plasma* (Jamieson, G. A., & Greenwald, T. J., Eds.) p 254, J. B. Lippincott Co., Philadelphia, PA.

Nathenson, S. G., Ewenstein, B. M., Uehara, H., Martinko, J. M., Coligan, J. E., & Kindt, T. J. (1979) in *Current Trends in Histocompatibility* (Ferrone, S., & Reisfeld, R.

A., Eds.) Plenum Press, New York (in press).

Natori, T., Tanigaki, N., Apella, E., & Pressman, D. (1975) Biochem. Biophys. Res. Commun. 65, 611.

Paul, W. E., & Benacerraf, B. (1977) Science 195, 1293.
 Rask, L., Lundstein, J. B., & Peterson, P. A. (1974) Nature (London) 249, 833.

Schwartz, B. D., Kato, K., Cullen, S. E., & Nathenson, S. G. (1973) Biochemistry 12, 2157.

Seegers, W. H. (1967) Blood Clotting Enzymology, Academic Press, New York.

Shearer, G. M., Rehn, T. G., & Garbarino, C. A. (1975) J. Exp. Med. 141, 1348.

Shimada, A., & Nathenson, S. G. (1969) *Biochemistry* 8, 4048.

Shreffler, D. C., & David, C. G. (1975) Adv. Immunol. 20, 125.

Silver, J., & Hood, L. E. (1974) Nature (London) 249, 764. Snell, G. D., Dausset, J., & Nathenson, S. G. (1976) Histocompatability, Academic Press, New York.

Strominger, J. L., et al. (1979) in *The Role of the Major Histocompatibility Complex in Immunobiology* (Benacerraf, B., & Dorf, M., Eds.) Marcel Dekker, New York (in press).

Vitetta, E. S., & Capra, J. D. (1978) Adv. Immunol. 26, 147.

Immunochemistry of Human Very Low Density Lipoproteins: Apolipoprotein C-III[†]

S. J. T. Mao, P. K. Bhatnagar, A. M. Gotto, Jr., and J. T. Sparrow*

ABSTRACT: Apolipoprotein C-III (apoC-III) is a major protein constituent of human plasma very low density lipoproteins (VLDL) and a minor constituent of high density lipoproteins (HDL). The apoprotein is a single polypeptide chain of 79 amino acids and occurs in several forms differing only in their content of sialic acid. In the present report a quantitative radioimmunoassay (RIA) has been developed in order to study the immunochemical properties of apoC-III. Two individual rabbit antisera were used. It was observed that the COOHterminal half (residues 41-79) of apoC-III containing 1 mol of sialic acid (apoC-III₁) contains the antigenic reactive region(s), whereas the NH2-terminal half of the molecule was unable to react with anti-apoC-III antibodies. When the positively charged lysines of apoC-III₁ were chemically modified by acetic anhydride, the immunoreactivity was decreased by ~40% as measured by RIA in one rabbit antise-

rum. However, there was no decrease of immunoreactivity in another rabbit antiserum, suggesting that with the former antibody several populations exist directed at different regions of the apoprotein molecule. In addition, the immunoreactivity of apoC-III₁-phospholipid complexes was investigated and was found to be indistinguishable from that of apoC-III₁. Since the conformation of the apoprotein has drastically changed upon the addition of dimyristoylphosphatidylcholine (DMPC), this finding indicates that the gross conformational change of apoC-III₁ does not affect the immunochemical properties and that the antigenic reactive sites are probably located at the surface in apoC-III₁-DMPC complexes. The immunoreactivity of apoC-III was also found to be approximately the same in HDL or VLDL as that of the delipidated apoHDL or apoVLDL. Thus, the antigenic sites of apoC-III must be fully exposed on the surface of the lipoproteins.

ApoC-III¹ is the most abundant of the C-proteins in human plasma very low density lipoproteins (VLDL). The protein exists in at least three forms which differ only in their content of sialic acid. The apoprotein contains 79 amino acids and has no cysteine, cystine, or isoleucine; the amino acid sequence (Figure 1) has been determined (Shulman et al., 1974; Brewer et al., 1974). The physicochemical and lipid-binding properties of the protein have been reviewed in great detail (Jackson et al., 1976; Morrisett et al., 1977; Bradley & Gotto, 1978; Smith

et al., 1978). A previous report from this laboratory (Sparrow et al., 1977) has shown that thrombin digestion of apoC-III₁ yielded two fragments (residues 1-40 and 41-79) of the protein by cleavage at Arg₄₀-Gly₄₁. Only the COOH-terminal half (residues 41–79) was able to interact with phospholipid. The immunochemical properties of apoC-III1 and the thrombin cleavage fragments have not been established. In the present study we report the immunoreactivity of the thrombin fragments and the role of lysine toward the immunochemical properties of apoC-III₁. Modification of the lysine residues was chosen because in the known antigenic structures of sperm whale myoglobin (Atassi, 1975) and of egg white lysozyme (Atassi & Lee, 1978) positively charged lysines were found in all of the antigenic reactive regions. Since the apoprotein displays a helical structure in isolated protein-phospholipid complexes (Morrisett et al., 1973), the role of conformation

[†]From the Department of Medicine, Baylor College of Medicine and The Methodist Hospital, Houston, Texas 77030. Received June 15, 1979. This material was developed by the Atherosclerosis, Lipids and Lipoproteins Section of the National Heart and Blood Vessel Research and Demonstration Center, Baylor College of Medicine, a grant-supported research project of the National Heart, Lung and Blood Institute, National Institutes of Health, Grant No. HL 17269, and by the American Heart Association.

^{*} Address correspondence to this author at the Baylor College of Medicine. J.T.S. is an Established Investigator of the American Heart Association.

[‡]Present address: Lipoprotein Research Unit, Department of Medicine, Mayo Clinic, Rochester, MN 55901.

¹ Abbreviations used: HDL, high density lipoproteins; apoA-I and apoA-II, apoprotein constituents of HDL; VLDL, very low density lipoproteins; apoC-III, an apolipoprotein constituent of VLDL; apoC-III₁, apoC-III containing 1 mol of sialic acid; DMPC, dimyristoylphosphatidylcholine; CD, circular dichroism; RIA, radioimmunoassay.