W. F. (1976) *J. Mol. Biol. 100*, 157–178.
Snyder, F., & Stephens, N. (1959) *Biochim. Biophys. Acta 34*, 244–245.
Somerharju, P., & Renkonen, O. (1979) *Biochim. Biophys. Acta 573*, 83–89.
Somerharju, P., Brotherus, J., Kahma, K., & Renkonen, O. (1977) *Biochim. Biophys. Acta 487*, 154–162.
Tjiong, H. B., & Debuch, H. (1978) *Hoppe-Seyler's Z. Physiol. Chem. 359*, 71–79.
Tjiong, H. B., Lepthin, J., & Debuch, H. (1978) *Hoppe-Seyler's Z. Physiol. Chem. 359*, 63–67.
Wherrett, J. R., & Huterer, S. (1972) *J. Biol. Chem. 247*, 4114–4120.
Wieland, O. (1974) in *Methods of Enzyme Analysis* (Bergmeyer, H. U., Ed.) 2nd English ed., pp 1404–1409, Academic Press, New York.

Primary Structure of Murine Major Histocompatibility Alloantigens: Amino Acid Sequence of the Cyanogen Bromide Fragment Ia (Positions 139–228) from the H-2K^b Molecule†

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

ABSTRACT: The complete amino acid sequence of the cyanogen bromide (CNBr) fragment Ia (CN-Ia) from the murine histocompatibility antigen H-2K^b has been obtained by using radiosequence methodology. This glycopeptide is the largest CNBr cleavage product of the H-2K^b molecule and extends from position 139 to position 228. The sequence determined for CN-Ia was 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-Asn-Gly-(Asn)-Ala-Thr-Leu-Leu-Arg-Thr-*Asp*-Ser-Pro-Lys-Ala-His-Val-Thr-His-His-Ser-Arg-Pro-*Asp*-*Asp*-Lys-Val-Thr-Leu-Arg-Cys-Trp-Ala-Leu-Gly-Phe-Tyr-Pro-Ala-*Asp*-Ile-Thr-Leu-Thr-Trp-Gln-Leu-Asn-Gly-Glu-Glu-Leu-Ile-Gln-*Asp*-Met. The data were obtained by analysis of fragments derived by thrombic, tryptic, chymotryptic, and V8 protease digestion of CN-Ia. A carbohydrate moiety is attached to Asn at position 176. Homology between this 90 amino acid stretch of H-2K^b and HLA-B7 [Orr, H. T., Lopez de Castro, J. A., Lancet, D., & Strominger, J. L. (1979) *Biochemistry 18*, 5711] is 68%, and differences are noted at positions 176, 177, and 178 which in the H-2 molecule are the attachment region for a second carbohydrate moiety. No carbohydrate was detected in this position for HLA-B7 [Orr, H. T. Lopez de Castro, J. A. Lancet, D., & Strominger, J. L. (1979) *Biochemistry 18*, 5711].

The classical histocompatibility antigens are the products of the polymorphic genes at the K, D, and L regions of the murine H-2 major histocompatibility complex (MHC)[1] [see reviews by Klein (1975, 1979), Snell et al. (1976), and Vitetta & Capra (1978)]. These products are integral cell surface glycoproteins containing ~340 amino acids (Schwartz et al., 1973), and they exist in the membrane in association with β_2-microglobulin, a polypeptide of molecular weight 12 000 (Rask et al., 1974; Silver & Hood, 1974; Natori et al., 1975). Although the primary function of these antigenic products remains obscure, recent data suggest that they play a role in the recognition of virally induced and other cell surface antigens, in a process termed "associative recognition" [see reviews by Paul & Benacerraf (1977), Shearer & Schmitt-Verhulst (1977), and Zinkernagel & Doherty (1979)].

Possibly related to their postulated role in cell recognition is the remarkable polymorphism of the K and D gene products, a property first discovered in early serological analysis of laboratory mouse strains [see reviews by Klein (1975) and Snell et al. (1976)] and more recently verified in wild mice (Zaleska-Rutczynska & Klein, 1977). An understanding of the molecular basis for the polymorphism as well as the function–structure relationships will be more approachable when the primary structures of the K, D, and L products have been elucidated.

Because of the difficulty in obtaining significant amounts of H-2 antigens in purified form, the most successful biochemical studies have utilized radiolabeled material obtained from cells cultured in the presence of radioactive amino acids. Primary structural studies of the MHC product coded for by the K gene of the *H-2^b* haplotype have also utilized radiochemical methods. Five major CNBr fragments of K^b were isolated, and they were provisionally aligned as follows: CN-IIIn (23 residues), CN-IIIa (29 residues), CN-Ib (86 residues), CN-Ia (90 residues), and CN-Ic (56 residues) (Ewenstein et al., 1978). The amino acid sequence of residues 1–173 has been reported (Uehara et al., 1980). The present report describes the sequence of CN-Ia (residues 139–228) and with data in the following paper (Martinko et al., 1980)

† 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 Institutes of Health, Bethesda, Maryland 20205 (T.J.K. and J.E.C.). *Received June 6, 1980.* These studies were supported in part by National Institutes of Health Grants AI-07289 and AI-10702 (S.G.N. and H.U.) and 5R32-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; H-2K$^b_{pap}$, H-2K^b glycoprotein derived by papain cleavage of the NP-40 solubilized H-2K^b molecule; β_2-m, β_2-microglobulin.

0006-2960/80/0419-6182$01.00/0

completes the primary structure of the papain fragment of the H-2K^b molecule.

Materials and Methods

Antisera, Radiolabeling, and Preparation of Cell Extract. The alloantiserum used to isolate the H-2K^b glycoprotein was prepared as previously described (Ewenstein et al., 1978; Uehara et al., 1980). Methods for the incorporation of radiolabeled amino acids into cells in culture and the preparation of a soluble cell extract have been described (Nairn et al., 1980; Uehara et al., 1980).

Isolation of the H-2K^b Molecule and of Its CNBr Fragments. The H-2K^b glycoprotein was isolated by immune precipitation from lectin affinity fractionated NP-40 cell extracts. Sephadex G-75 column chromatography was used to separate the noncovalently associated β_2-m (Ewenstein et al., 1978). The H-2K^b papain fragment (H-2K$^b_{pap}$, M_r 37 000) was isolated by papain digestion of H-2K^b immune precipitates derived from the NP-40 solubilized H-2K^b molecule as described previously (Ewenstein et al., 1976; Nairn et al., 1980). Reduction and alkylation of H-2K$^b_{pap}$ were performed as described (Brown et al., 1974). The CNBr fragments of the H-2K^b molecule were obtained and purified as described by Ewenstein et al. (1978).

Automated Amino Acid Sequence Analysis. Automated Edman degradation was performed with a Beckman Sequenator (Model 890-C) by using Beckman Peptide Program No. 102974. The detection of radioactivity, conversion of thiazolinone amino acids to phenylthiohydantoin (PTH) derivatives, and identification of PTH amino acids were performed as described previously (Coligan et al., 1979; Uehara et al., 1980).

Thrombin Digestion. Thrombin (B grade, Lot No. 703053, Calbiochem, La Jolla, CA) was prepared as described previously (Uehara et al., 1980). CN-Ia samples labeled with radioactive amino acids and containing 5–7 mg of carrier cytochrome *c* and bovine serum albumin were dissolved in 0.7–1.0 mL of 0.1 M NH_4HCO_3, pH 8.55, and incubated at 4 °C for 10–13 h with 0.09–0.12 mL of thrombin solution (activity, ΔA_{247}/(min mL), = 1.6 at 23 °C) (Uehara et al., 1980). The reaction was halted by the addition of 0.1 mL of 88% HCOOH or 1–1.2 mL of CH_3COOH, and the mixture was applied directly on a column of Sephadex G-50 (Superfine) equilibrated with either 1 M HCOOH or 50% (v/v) CH_3COOH.

Chymotrypsin Digestion. CN-Ia (containing 2 mg of carrier cytochrome *c*) was dissolved in 1 mL of 0.1 M NH_4HCO_3, pH 8.55, and incubated at 37 °C for 60 min with α-chymotrypsin (Worthington Biochemical Corp., Freehold, N.J.) at a substrate to enzyme weight ratio of 100:2. The reaction was halted by addition of 0.5 mL of CH_3COOH and applied directly on a Spherix ion-exchange column and developed by a linear pH and pyridine gradient as described previously (Brown et al., 1974).

Trypsin digestion and Staphylococcus V8 protease digestion were performed as described previously (Brown et al., 1974; Uehara et al., 1980).

Results

Isolation and Sequence Determination of Peptides Obtained by Limited Thrombin Digestion of CN-Ia. CN-Ia, the largest fragment obtained from CNBr cleaved H-2K^b, is a glycopeptide of 90 amino acids extending from residue 139 to residue 228. Thrombin digestion of CN-Ia was carried out under limited reaction conditions to obtain subfragments suitable for sequence analysis. The thrombic digest was fractionated by gel filtration on a Sephadex G-50 column equilibrated either with 1 M HCOOH or with 50% CH_3COOH (Figure 1). Gel filtration in 1 M HCOOH (Figure 1A) separated two major peptides (θ1 and θ2). An additional peptide (θ3) was isolated if the digestion mixture was chromatographed in 50% CH_3COOH (Figure 1B).

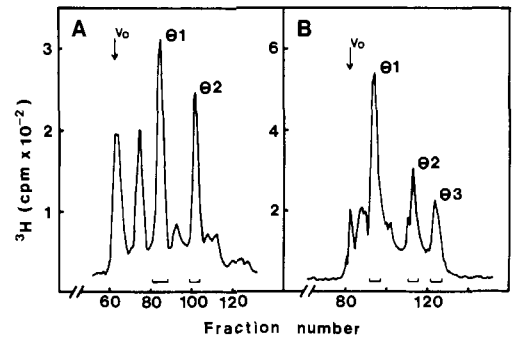

FIGURE 1: Gel filtration of thrombic peptides of CN-Ia on a Sephadex G-50 column. (A) A sample labeled with [^{3}H]SIGHTVW was chromatographed on a Sephadex G-50 (Superfine) column (0.9 × 170 cm) equilibrated with 1 M HCOOH. The flow rate was 2.2 mL/h, and 0.75-mL fractions were collected. The peptides were detected by sampling each fraction for radioactivity; pools were made as indicated. V_0 indicates the void volume of the column. (B) A sample labeled with [^{3}H]FLKYTVR was chromatographed on a Sephadex G-50 (Superfine) column (1.5 × 150 cm) equilibrated with 50% CH_3COOH. The flow rate was 6.2 mL/h, and 1.3-mL fractions were collected.

Each thrombic peptide had a unique NH_2-terminal sequence. Partial amino acid sequence studies of the three peptides enabled us to localize θ1 to the NH_2 terminus of CN-Ia. Only θ3 contained [^{3}H]Met and therefore was thought to include the COOH-terminal residue of CN-Ia. Both θ1 and θ3 contained [^{35}S]Cys whereas this amino acid was absent in θ2. In addition, only θ1 contained [^{3}H]fucose and hence carried the carbohydrate moiety. Because θ1 and θ2 could be obtained in relatively high yield and better purity in the 1 M HCOOH system, this system was chosen for the isolation of θ1 and θ2.

The sequence determination of θ3 was initially carried out by an NH_2-terminal sequence analysis of θ3 isolated by gel filtration in 50% CH_3COOH. However, following exposure to 50% CH_3COOH during the course of purification, the NH_2-terminal of θ3 (Cys) appeared to be partially blocked and yielded poor recoveries when subjected to sequence analysis. Consequently the majority of assignments after residue 15 in θ3 were determined by sequence analysis of a tryptic peptide TC2 isolated from papain-treated H-2K^b (H-2K$^b_{pap}$) as described below. The radioactivity recovered at each step in the sequence analyses of CN-Ia, θ2, and θ3 is shown in Figure 2A–C.

The partial amino acid sequence analysis of thrombic peptide θ1 (Table I) confirmed many of the assignments made from the NH_2-terminal sequence analysis of intact CN-Ia. The amino acid assignments in θ2, which is a polypeptide of 21 amino acids, were made by NH_2-terminal sequence analysis of the peptide. The COOH-terminal residue was found to be Arg, consistent with the known proteolytic specificity of thrombin.

Isolation of Tryptic Peptides from CN-Ia. In order to determine the COOH-terminal sequence of θ1 and to confirm the assignments made by the NH_2-terminal sequence analysis of CN-Ia, θ1 was digested by trypsin, and the peptides were purified by Spherix cation-exchange column chromatography (Figure 3). The breakthrough peptide (T8), which was further purified by gel filtration on a Sephadex G-50 column

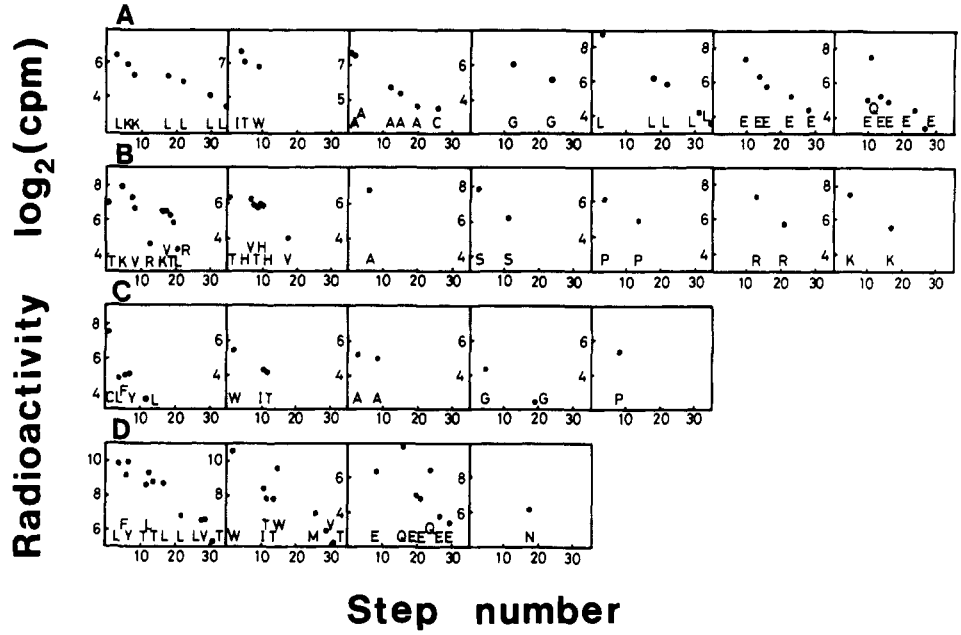

FIGURE 2: Radioactivity [as $\log_2$ (cpm)] recovered in each step of amino acid sequence analysis plotted against step number. The amino acid residue assigned at each position is indicated in single letter amino acid code. (A) Sequence analysis of CN-Ia. (B) Sequence analysis of Ia-θ2. (C) Sequence analysis of Ia-θ3. (D) Sequence analysis of TC-2.

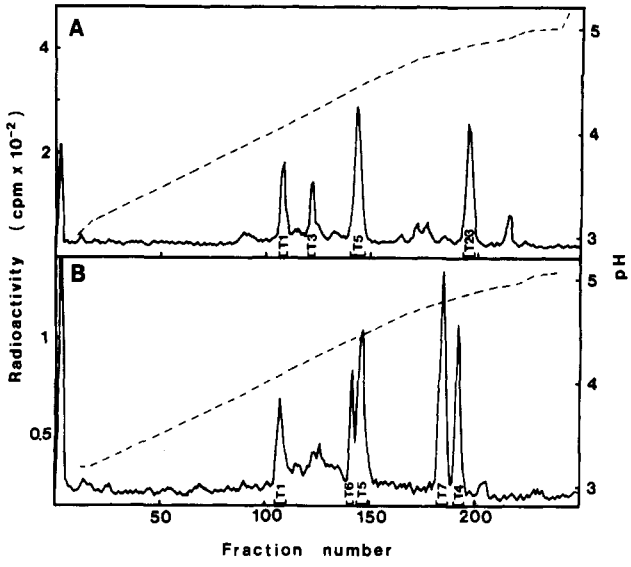

FIGURE 3: Ion-exchange column chromatography of tryptic peptides of Ia-θ1 labeled with [^{3}H]SIGHTVW (A) and [^{3}H]FLKYR/[^{35}S]C (B). Details of chromatographic conditions were described by Brown et al. (1974). The products not bound to the column (fractions 2–5) were further purified by gel filtration on a column of Sephadex G-50 (0.9 × 170 cm) equilibrated with 1 M HCOOH.

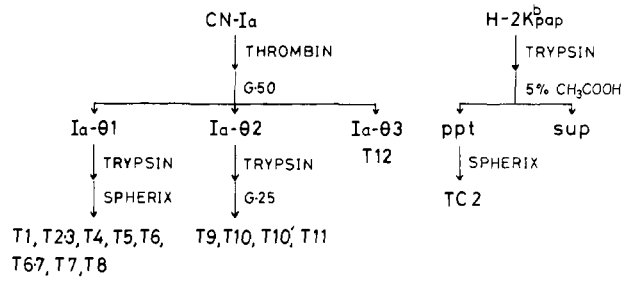

FIGURE 4: Flow diagram for purification of tryptic peptides in CN-Ia. T10′ is a product of inappropriate cleavage at a His residue. T12 is identical to Ia-θ3. TC-2 includes Ia-θ3 and the NH_2-terminal 15 residues of CN-Ic. The sequence analysis for each peptide is shown in Figure 6 and documented in Table I.

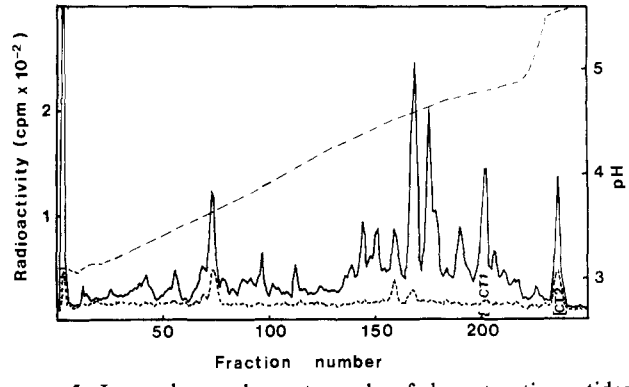

FIGURE 5: Ion-exchange chromatography of chymotryptic peptides of CN-Ia radiolabeled in [^{3}H]WPLRS/[^{35}S]C. CT1 and CT2 were pooled as shown by the bar, and their sequences were determined.

in 1 M HCOOH, was found to be the glycopeptide since it contained ^{3}[H]fucose. Amino-terminal sequence analysis of CN-Ia and sequence analysis of a peptide obtained by V8 cleavage of CN-Ia, as described later, allowed us to localize peptide T8 to the COOH terminus of θ1. The assignments at positions 7, 25, 27, 29, 33, 36, 37, 38, 40, 41, 42, and 43 of CN-Ia were determined by sequence analysis of these tryptic peptides and V8 peptides as described later.

The peptides T9, T10, T11, and T10′ were isolated from tryptic digests of θ2 by gel filtration on a Sephadex G-25 column in 1 M HCOOH (data not shown). The isolation of these peptides is summarized in Figure 4 and their amino acid sequence shown in Figure 6.

Alignment of the Thrombic Peptides. The thrombic peptides were aligned by the isolation of overlapping peptides and by the specific labeling of θ3 with [^{3}H]Met. CN-Ia labeled by [^{3}H]FLKYTVR was cleaved by V8 protease, and one large peptide (V8-1) was purified by gel filtration on a Sephadex G-50 column. The amino acid sequence of this peptide was determined to be -LRRYLK-----LLR----K--V--, thereby aligning θ1 and θ2. Peptides obtained by chymotryptic cleavage of [^{3}H]WPLRS/[^{35}S]C-labeled CN-Ia were fractionated by ion-exchange column chromatography (Figure 5). The sequence of CT1 was found to be R--SP- and that of

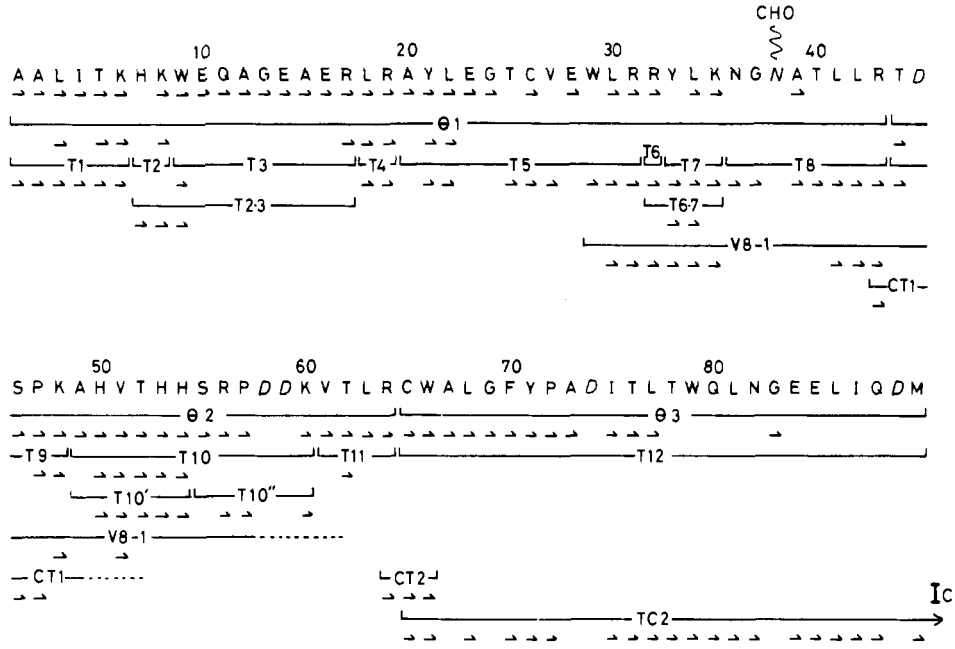

FIGURE 6: Summary of the amino acid sequence information for CN-Ia. Numbers above sequence refer to residue positions in fragment Ia. The first line of arrows depicts residues identified by NH_2-terminal sequence analysis of CN-Ia. Other arrows depict the sequences for thrombic peptides (θ), tryptic peptides (T, TC), staphylococcus protease V8 peptides (V8), and chymotryptic peptides (CT).

CT2 was determined to be RCW. Although the isolation of CT2 does not eliminate the possible presence of small peptides between $\theta 2$ and $\theta 3$, the thrombic peptides were aligned as $\theta 1$-$\theta 2$-$\theta 3$. Confirmation of this alignment is the finding that the amino acid sequence of CN-Ia shows a high degree of homology with the corresponding HLA-B7 sequence (Orr et al., 1979).

Isolation of Tryptic Peptide (TC-2) from H-2K$^b_{pap}$. Because of the difficulty in determining the sequence of $\theta 3$ by the procedure described above, we decided to isolate a tryptic peptide containing $\theta 3$ from the whole H-2K$^b_{pap}$ molecule.

Reduced and alkylated H-2K$^b_{pap}$ was digested by trypsin, and sufficient 100% CH_3COOH was subsequently added to the reaction mixture to make a final concentration of 5% (v/v). The peptides precipitated by this procedure (TC, tryptic core) were collected by centrifugation (10^4 g·min), washed with 5% CH_3COOH, dissolved in 65% CH_3COOH, and applied on a Spherix ion-exchange column (Figure 4). The fraction which was excluded from the column was found to contain only one peptide (TC-2). Amino acid sequence analysis of TC-2 proved that this peptide was the overlap peptide between CN-Ia and CN-Ic and contained all the residues of CN-Ia-$\theta 3$ and the NH_2-terminal 15 residues of CN-Ic. Extensive sequence studies of TC-2 were then performed to determine the sequence of the region corresponding to $\theta 3$ (Figure 2D).

Summary of Amino Acid Sequence Data for CN-Ia. The complete amino acid sequence of CN-Ia determined from thrombic, tryptic, chymotryptic and V8 protease peptides is summarized in Figure 6. Table I gives the documentation of the assignments. Asp is assigned to any position for which there is no positive assignment after definitive sequence determination has been carried out for all other amino acids. Positions 45, 58, 59, 74, and 89 were assigned as Asp.

Tryptic peptide T8 (residues 36–43) was specifically labeled by [^{3}H]fucose and is one of the two carbohydrate-containing peptides of H-2K^b. Although Asn was assigned at position 36 of CN-Ia by automated sequence analysis of T8, position 38 was also tentatively assigned as Asn. Thus the sequence at positions 38–40 is Asn-Ala-Thr which satisfies the recognition sequence for glycosylation (Marshall, 1972), and residue 38 is presumed to be the carbohydrate attachment site.

Discussion

The amino acid sequence of the CNBr fragment Ia (CN-Ia) obtained from the H-2K^b glycoprotein has been determined. CN-Ia is a polypeptide of 90 amino acid residues containing one of the two carbohydrate moieties found in H-2K^b. The amino acid sequence determination of CN-Ia was completed by sequence analysis of subfragments derived by thrombic, tryptic, chymotryptic, and V8 protease cleavage of CN-Ia as well as from tryptic peptides obtained by digestion of intact H-2K$^b_{pap}$. The determination of the amino acid sequence of this fragment, which extends from residue 139 to residue 228, when combined with the data of the CN-Ic fragment documented in the following paper (Martinko et al., 1980), completes the primary structure of the papain fragment of the H-2K^b molecule.

Thrombin, a protease of extremely narrow specificity (Seegers, 1967), was used to obtain large subfragments of CN-Ia. However, a carefully chosen reaction temperature and the presence of the appropriate nonradiolabeled carrier were needed to assure reproducibility and high efficiency of the desired cleavages. Thus, while cleavage at 4 °C occurred at two peptide bonds preferentially producing three major peptides, additional cleavages localized within the NH_2-terminal 43 residues were observed at higher reaction temperatures (23 or 37 °C).

The nature of the protein carrier greatly affected the yield of the thrombic peptides. Thus, a mixture of bovine serum albumin and cytochrome *c* was found to be an appropriate carrier whereas reduced and alkylated cytochrome *c* gave poor results, presumably because of its poor solubility under the conditions employed.

Ia-$\theta 3$ was not isolated by gel filtration in 1 M HCOOH because of the low solubility of this peptide in this solvent, but purification of Ia-$\theta 3$ could be achieved by the use of gel filtration in 50% CH_3COOH. However, the use of 50% CH_3COOH appeared to cause an NH_2-terminal modification of the peptide since an extremely low initial yield was observed on sequence analysis. Although no further studies to inves-

Table I: Identification of Amino Acids in CN-Ia

		NH_2-terminal		thrombin			trypsin		
fragment no.	assigned residue[a]	label group[b]	ID	peptide	label group[b]	ID	peptide	label group[b]	ID
1	A	III		θ1			T1	Id	LC
2	A	III						Id	LC
3	L	Ia, S	LC, TLC		Ic, S	LC		Ib, Id	LC
4	I	IIa	TLC					IIa	LC
5	T	IIa	TLC		Ic	LC		IIa	LC
6	K	Ia	LC, TLC		Ic	LC		Id	LC
7	H						T2, T2·3	IIa	LC
8	K	Ia	LC, TLC		Ic	LC		IIb	LC
9	W	IIa	TLC				T3	IIa	LC
10	E	IVa, IVb	LC						
11	Q	IVb	LC						
12	A	III							
13	G	Va, Vb							
14	E	IVa, IVb	LC						
15	A	III							
16	E	IVa, IVb	LC						
17	R	S			Ic	LC			
18	L	Ia, S	LC		Ic, S	LC	T4	Ib	LC
19	R	S			Ic	LC		Ib	
20	A	III					T5		
21	Y	S			Ic	LC		Ib	LC
22	L	Ia, S	LC		Ic	LC		Ib	LC
23	E	IVa, IVb	LC						
24	G	Va, Vb							
25	T							IIa	LC
26	C	III, S						Ib	
27	V							IIa	LC
28	E	IVa, IVb	LC						
29	W							IIa	LC
30	L	IIa, S	LC					Ib	LC
31	R	S						Ib	LC
32	R	S					T6, T6·7	Ib	LC
33	Y						T7, T7·8	Ib	LC
34	L	S						Ib	LC
35	K	S						Ib	LC
36	N						T8	S	LC
37	G							Va	LC
38	(N)								
39	A	S						S	LC
40	T							IIa	LC
41	L							Ib	LC
42	L							Ib	LC
43	R							Ib	
44	T			θ2	IIa	LC	T9		
45	(D)								
46	S				Va, Vb	LC			
47	P				IVa, S	LC		Id	LC
48	K				S			Id	LC
49	A				S		T10, T10′		
50	H				IIa	LC		Ia	LC
51	V				IIa	LC		Ia	LC
52	T				Ic, IIa, S	LC		IIa	LC
53	H				IIa	LC		IIa	
54	H				IIa	LC		IIa	
55	S				Va, Vb	LC	T10		
56	R				S2			Id	LC
57	P				IVa, S	LC		Id	LC
58	(D)								
59	(D)								
60	K				S			Id	LC
61	V				IIa	LC	T11		
62	T				Ic, S	LC			
63	L				Ic	LC			
64	R				S2				
65	C			θ3	Ib, S		TC-2	S	
66	W				IIb	LC		IIb	LC
67	A				S	LC			
68	L				Ib	LC		Ic	LC
69	G				Va, Vb	LC			
70	F				Ib	LC		Ic	LC
71	Y				Ib	LC		Ic	LC
72	P				S	LC		IVb	
73	A				S	LC			
74	(D)								
75	I				IIb	LC		IIc	LC

Table I (*Continued*)

fragment no.	assigned residue[a]	NH_2-terminal label group[b]	NH_2-terminal ID	thrombin peptide	thrombin label group[b]	thrombin ID	trypsin peptide	trypsin label group[b]	trypsin ID
76	T				IIb	LC		Ic, IIc	LC
77	L				Ib	LC		Ic	LC
78	T							Ic, IIc	LC
79	W							IIc	LC
80	Q							IVb	LC
81	L							Ic	LC
82	N							S	LC
83	G				Va, Vb				
84	E							IVb	LC
85	E							IVb	LC
86	L							Ic	LC
87	I							IIc	LC
88	Q							IVb	LC
89	(D)								
90	M							IIc, S	

[a] The standard one-letter symbols for amino acids (IUPAC-IUB Commission on Biochemical Nomenclature, 1968) 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 Ia-Vb. Ia, [^{3}H]FLKY; Ib, [^{3}H]FLKYR/[^{35}S]C; Ic, [^{3}H]FLKYTVR; Id, [^{3}H]FALKPYR/[^{35}S]C; IIa, [^{3}H]SIGHTVW; IIb, [^{3}H]HITVW; IIc, [^{3}H]HITVW/[^{35}S]M; III, [^{3}H]AC; IVa, [^{3}H]EP; IVb, [^{3}H]QEP; Va, [^{3}H]GS; Vb, [^{3}H]SG. All radiolabeled PTH amino acids in multilabeled preparations and many in single radiolabeled preparations were identified by high-pressure liquid chromatography (LC) and/or thin-layer chromatography (TLC) as indicated.

tigate the modification reaction were performed, one possible reaction is the conversion of N-terminal *S*-(carboxamidomethyl)cysteine to a thiazanecarboxamide residue under acid conditions. In the case of NH_2-terminal carboxymethylcysteine, conversion to a thiazane carboxyl residue has been reported (Smyth & Utsumi, 1967).

A tryptic peptide, TC-2, isolated from digests of H-2K$^b_{pap}$ was used in analysis of COOH-terminal portion of CN-Ia. TC-2 contains the peptide Ia-θ3 plus an additional 15 amino acid residues. The procedure used for the purification of TC-2 involves only a short exposure to a high acid concentration in order to avoid the potential cyclization reaction at the NH_2-terminal residue of TC-2 as described above. The amino acid sequence analysis of TC-2 permitted assignment at every position in Ia-θ3 and confirmed the alignment of CN-Ia and CN-Ic.

Of the 90 amino acid residues in CN-Ia, 84 residues were definitely assigned by sequence analysis of the radiolabeled peptides examined (Table I). As described under Results, position 45, 58, 59, 74, and 89 are tentatively assigned as Asp. Among five Asp residues assigned tentatively, three at positions 45, 58, and 89 are homologous with the corresponding HLA-B7 sequence. Furthermore, the Asp at 74 in H-2K^b is assigned as Glu in HLA-B7, which is a highly conservative change. The Asn assignment at position 38 and the conclusion that this is the carbohydrate attachment site are based on the fact that tryptic peptide T8 (residues 36–43) is a carbohydrate-containing peptide and that position 38 is the only one that contains a potential glycosylation sequence (Asn-X-Thr/Ser) (Marshall, 1972) and the only position in T8 at which an amino acid residue was not identified during sequence analysis. Glycosylated Asn residues are not extracted from the spinning cup during automated sequence analysis and hence are not detected. Further experiments to confirm the assignment of an Asn residue at position 38 are in progress.

The comparison of the amino acid sequence between CN-Ia and the corresponding HLA-B7 (Orr et al., 1979) positions (138–229) shows 68% homology. Most of the interchanges require only single-base substitutions. Interestingly, three contiguous interchanges are observed at positions 38, 39, and 40 in CN-Ia, which correspond to 176, 177, and 178 in the intact K^b and HLA-B7 molecules. Position 176 is thought to be a carbohydrate chain attachment site in the H-2K^b molecule. HLA-B7 does not contain a carbohydrate moiety in this region (Orr et al., 1979). This is a major structural difference between the two molecules, a feature which is further discussed in the following paper.

The data described in this report in combination with the data on the amino acid sequence of CN-Ic, which is described in the following paper, complete the amino acid sequence of H-2K$^b_{pap}$. A detailed discussion of the overall structural features and comparisons with other molecules are made in the following paper (Martinko et al., 1980).

Acknowledgments

We acknowledge the expert technical assistance of Mark Flocco, Diane McGovern, Selma Schumacher, Pat Spinella, and Robert Valas and the secretarial assistance of Catherine Whelan. We also thank Drs. Roderick Nairn and Karen Yamaga for fruitful discussions of the data and the manuscript.

References

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., Martinko, J. M., & Nathenson, S. G. (1979) *Mol. Immunol. 16*, 3.

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.

IUPAC–IUB Commission on Biochemical Nomenclature (1968) *J. Biol. Chem. 243*, 3557.

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

Klein, J. (1979) *Science (Washington, D.C.) 203*, 516.

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

Martinko, J. M., Uehara, H., Ewenstein, B. M., Kindt, T. J., Coligan, J. E., & Nathenson, S. G. (1980) *Biochemistry* (following paper in this issue).

Nairn, R., Nathenson, S. G., & Coligan, J. E. (1980) *Eur. J. Immunol.* (in press).

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

Orr, H. T., Lopez de Castro, J. A., Lancet, D., & Strominger, J. L. (1979) *Biochemistry 18*, 5711.
Paul, W. E., & Benacerraf, B. (1977) *Science (Washington, D.C.) 195*, 1293.
Rask, L., Lindblom, 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) in *Blood Clotting Enzymology*, Academic Press, New York.
Shearer, G. M., & Schmitt-Verhulst, A. (1977) *Adv. Immunol. 25*, 55.
Silver, J., & Hood, L. E. (1974) *Nature (London) 249*, 764.
Smyth, D. S., & Utsumi, S. (1967) *Nature (London) 216*, 332.
Snell, J. D., Dausset, J., & Nathenson, S. G. (1976) in *Histocompatibility*, Academic Press, New York.
Uehara, H., Ewenstein, B. M., Martinko, J. M., Nathenson, S. G., Coligan, J. E., & Kindt, T. J. (1980) *Biochemistry 19*, 306.
Vitetta, E. S., & Capra, J. D. (1978) *Adv. Immunol. 26*, 147.
Zaleska-Rutczynska, Z., & Klein, J. (1977) *J. Immunol. 119*, 1903.
Zinkernagel, R. M., & Doherty, P. C. (1979) *Adv. Immunol. 27*, 51.

Primary Structure of Murine Major Histocompatibility Complex Alloantigens: Completion of the Sequence of the Amino-Terminal 284 Residues of H-2K^{b}†

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

ABSTRACT: The primary structure of the COOH-terminal cyanogen bromide (CNBr) cleavage fragment Ic (CN-Ic) of the extracellular portion of the murine histocompatibility antigen H-2K^{b} has been completed. CN-Ic contains a site of papain cleavage which has been utilized for solubilizing H-2K^{b} by cleaving off the membrane integrating portion of the molecule. The amino acid sequence of CN-Ic has been determined by using peptides recovered after trypsin digestion of CN-Ic before and after blockage of lysine groups with citraconic anhydride. Overlapping sequences for the tryptic fragments were obtained by amino-terminal sequence analysis. The sequence of fragment CN-Ic, which spans residues 229–284 in H-2K^{b}, is as follows: Glu-Leu-Val-Glu-Thr-Arg-Pro-Ala-Gly-Asp-Gly-Thr-Phe-Gln-Lys-Trp-Ala-Ser-Val-Val-Pro-Leu-Gly-Lys-Glu-Gln-Tyr-Tyr-Thr-Cys-His-Val-Tyr-Gln-Gln-Gly-Leu-Pro-Gln-Pro-Leu-Thr-Leu-Arg-Trp-Asp-Glu-Pro-Pro-Ser-Thr-Val-Ser-Asn-Met. This amino acid sequence determination completes the primary structure of the amino terminal 284 residues of H-2K^{b}, that portion of this histocompatibility antigen which is external to the cell membrane and which contains antigenic determinants. It was also possible to identify Val-281 as a papain cleavage site within CN-Ic. The completed structure was analyzed solely by radiochemical methods. The structure obtained for H-2K^{b} is 71% homologous to the reported structure of HLA-B7, a human homologue.

Radiochemical microsequencing techniques have been applied to the murine H-2 alloantigen H-2K^{b} [1] as an approach to determining the primary structure of proteins available in amounts too small for classical sequencing analysis. Determination of the sequence of the first 173 residues was accomplished in a series of studies on the CNBr fragments, CN-IIIn, CN-IIIa, and CN-Ib (Coligan et al., 1978, 1979; Uehara et al., 1980a). The sequence of CN-Ia (residues 139–228) is described in the preceding paper (Uehara et al., 1980b). The data for determination of CN-Ic (residues 229–284) are presented in the present paper, permitting assignment of the NH_2-terminal 284 amino acids (H-2K$^{b}_{284}$) which is approximately equivalent in size to the H-2K^{b} papain-solubilized fragment (H-2K$^{b}_{pap}$).

Materials and Methods

Preparation of Radiolabeled H-2K^{b} Antigen. The H-2K^{b} glycoprotein was isolated from EL-4.BU cells which had been cultured in the presence of various radiolabeled amino acids. After solubilization of the cell membrane in Nonidet P-40, the glycoprotein fraction was isolated by *Lens culinaris* hemagglutinin affinity chromatography, indirect immune precipitation, papain digestion (if indicated), and G-75 column chromatography as previously described (Ewenstein et al., 1978).

Preparation of CN-Ic. The disulfide-bonded complex of the major CNBr fragments of H-2K^{b}, which includes CN-Ia, CN-Ib, and CN-Ic, was purified by Sephadex G-100 chromatography. Reduction and alkylation of this material were done with 0.1 M dithiothreitol under N_2 at 37 °C for 2 h and 0.25 M iodoacetamide for 15 min at ambient temperature. CN-Ic was then isolated by Sephadex G-100 column chromatography. A more detailed description of these purification methods was previously published (Ewenstein et al., 1978).

Citraconylation of CN-Ic. Citraconylation of CN-Ic was accomplished by using a modification of previously described methods (Dixon & Perham, 1969). Radiolabeled CN-Ic and 2 mg of carrier bovine serum albumin were modified by re-

† From the Department of Microbiology and Immunology and the Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461 (J.M.M., H.U., B.M.E., and S.G.N.), and the Laboratory of Immunogenetics, National Institutes of Health, Bethesda, Maryland 20205 (T.J.K. and J.E.C.). *Received June 6, 1980.* These studies supported in part by National Institutes of Health Grants 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: H-2K$^{b}_{284}$, H-2K^{b} glycopeptide comprising the NH_2-terminal 284 amino acid residues of intact molecule; MHC, major histocompatibility complex; HLA-B7$_{pap}$, HLA-B7 isolated after papain cleavage of membrane containing intact HLA-B7.

0006-2960/80/0419-6188$01.00/0