Makabe, O., Nakamura, M., & Umezawa, S. (1975) *J. Antibiot. 28*, 492–495.
Manor, P. C., Saenger, W., Davies, D. B., Jankowski, K., & Rabczenko, A. (1974) *Biochim. Biophys. Acta 340*, 472–483.
Matsuda, A., Tezuka, M., Niizuma, K., Sugiyama, E., & Ueda, T. (1978) *Tetrahedron 34*, 2633–2637.
Murray-Rust, P., & Motherwell, S. (1978) *Acta Crystallogr., Sect. B 34*, 2534–2546.
Parthasarathy, R., Ohrt, J. M., & Chheda, G. B. (1977) *Biochemistry 16*, 4999–5008.
Pless, R., Dudycz, L., Stolarski, R., & Shugar, D. (1978) *Z. Naturforsch., C: Biosci. 33*, 902–907.
Plochocka, D., Rabczenko, A., & Davies, D. B. (1977) *Biochim. Biophys. Acta 476*, 1–15.
Raleigh, J. A., & Blackburn, B. J. (1978) *Biochem. Biophys. Res. Commun. 83*, 1061–1066.
Remin, M., & Shugar, D. (1973) *J. Am. Chem. Soc. 95*, 8146–8156.
Sharples, T. R., & Wolfenden, R. (1967) *Methods Enzymol. 12A*, 126–131.
Simon, L. N., Bauer, R. J., Tolman, R. L., & Robins, R. K. (1970) *Biochemistry 9*, 573–577.
Stolarski, R., Dudycz, L., & Shugar, D. (1980) *Eur. J. Biochem. 108*, 111–121.
Uesugi, S., & Ikehara, M. (1978) *Chem. Pharm. Bull. 26*, 3040–3049.
Watenpaugh, K., Dow, J., Jensen, L. H., & Furberg, S. (1968) *Science (Washington, D.C.) 159*, 206–207.
Westhof, E., & Sundaralingam, M. (1980) *J. Am. Chem. Soc. 102*, 1493–1500.
Zemlicka, J. (1975) *J. Am. Chem. Soc. 97*, 5896–5903.

Amino Acid Sequence of Residues 1–98 of the H-2K^d Murine Major Histocompatibility Alloantigen: Comparison with H-2K^b and H-2D^b Reveals Extensive Localized Differences†

Edward S. Kimball, Stanley G. Nathenson, and John E. Coligan*

ABSTRACT: The complete primary structure of the amino-terminal 98 residues of the murine histocompatibility antigen H-2K^d has been determined by using radiosequencing methodology. Cyanogen bromide cleavage of the detergent-solubilized molecule yielded several peptides, two of which span residues 1–52 and 53–98. The latter peptide bears one of the carbohydrate moieties of the H-2K^d molecule. The complete sequence of the amino-terminal fragment (residues 1–52) was accomplished by analyses of peptides derived from cleavage by staphylococcal V8 protease. Analyses of peptides obtained from tryptic digests and NH_2-terminal sequence data were used to confirm the alignment of the V8 peptides. The complete sequence of the glycopeptide spanning residues 53–98 was achieved by NH_2-terminal sequence analyses and from fragments derived by thrombic cleavage. Alignment of the two CNBr fragments was deduced from a tryptic overlap peptide derived from the whole molecule. Positive identification was possible for all amino acids except Asp and the C-terminal Met (position 98). The sequence obtained in the study is Gly-Pro-His-Ser-Leu-Arg-Tyr-Phe-Val-Thr-Ala-Val-Ser-Arg-Pro-Gly-Leu-Gly-Glu-Pro-Arg-Phe-Ile-Ala-Val-Gly-Tyr-Val-*Asp*-*Asp*-Thr-Gln-Phe-Val-Arg-Phe-*Asp*-Ser-*Asp*-Ala-*Asp*-Asn-Ala-Arg-Phe-Glu-Pro-Arg-Ala-Pro-Trp-Met-Glu-Gln-Glu-Gly-Pro-Glu-Tyr-Trp-Glu-Glu-Gln-Thr-Gln-Arg-Val-Lys-Ser-*Asp*-Glu-Gln-Trp-Phe-Arg-Val-Ser-Thr-Arg-Thr-Ala-Gln-Arg-Tyr-Tyr-Asn-Gln-Ser-Lys-Gly-Gly-Ser-His-Thr-Phe-Gln-Arg-*Met*. These data allow the first extensive comparison of the primary structure of two allelic histocompatibility gene products (H-2K^d and H-2K^b). This comparison revealed one 22-residue-long region of amino acid sequence which contained 12 of the 23 amino acid differences noted between these molecules. Eleven of the remaining differences were not distributed randomly but localized to four discrete regions. No amino acid sequence distinctive of H-2K gene products was revealed when the NH_2-terminal 98 residues of the H-2K^d, H-2K^b, and H-2D^b amino acid sequences were compared.

One of the most striking aspects of the classical transplantation antigens of the murine histocompatibility (H-2) system is their high level of genetic polymorphism, as demonstrated by serological and cellular assays (reviewed by Klein, 1975; Snell et al., 1976). A systematic study of those H-2 antigens (H-2K, H-2D, and H-2L) has been undertaken in order to discern on a molecular level the basis for antigenic polymorphism and to ascertain the extent of evolutionary relationships that might prevail among H-2 antigens.

These H-2 antigens have been shown to be integral membrane glycoproteins (Shimada & Nathenson, 1969; Schwartz et al., 1973) that have approximate molecular weights of 45 000 and exist in noncovalent association with β_2-microglobulin (β_2m),[1] a nonintegrated protein (Rask et al., 1974; Nakamuro et al., 1973; Cresswell et al., 1973). The recent

† From the Laboratory of Immunogenetics, Building 8, Room 100, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205 (E.S.K. and J.E.C.), and The Albert Einstein College of Medicine, Bronx, New York 10461 (S.G.N.). *Received September 29, 1980.* Part of this study was supported by Grants AI-07289 and AI-10702 from the National Institutes of Health (S.G.N.).

[1] Abbreviations used: β_2m, β_2-microglobulin; HPLC, high-performance liquid chromatography; Gdn·HCl, guanidine hydrochloride; F_3Ac-OH, trifluoroacetic acid.

0006-2960/81/0420-3301$01.25/0

elucidation of the complete amino acid sequence of the extracellular portion of the H-2K^b molecule (Uehara et al., 1980; Martinko et al., 1980) advances our knowledge of these antigens still further and provides data for comparison of primary structures of all such antigens.

This report extends preliminary structural data obtained for the H-2K^d molecule (Kimball et al., 1980) and allows the first extensive primary structural comparison of allelic histocompatibility antigen gene products. Comparison of the first 98 residues of H-2K^d to those of the H-2K^b molecule revealed an overall homology of 77%; comparison to the H-2D^b gene product showed 75% homology. However, a region of greater than average variability exists between residues 62 and 83 which shows 45% and 50% homology to the H-2K^b and H-2D^b molecules, respectively. Other smaller clusters of sequence variability are also present outside this particular region. No primary structure distinctive of H-2K region gene products is evident.

Materials and Methods

Cell Line and Preparation of Radiolabeled Antigen. The mouse tumor cell line C14 (H-2^d) was used as the source of H-2K^d (Nairn et al., 1980a). The incorporation of radiolabeled amino acids was carried out by modifications (Nairn et al., 1980a) of a previously published procedure (Brown et al., 1974).

Antisera. The anti-H-2K^d alloantiserum was prepared as described by Snell (1968) in A/J mice (K^k, D^d) by injection with Meth A (K^d, D^d) tumor cells.

Isolation of H-2K^d Glycoprotein. The H-2K^d glycoprotein was isolated by immunoprecipitation from glycoprotein pools which had been obtained by lentil lectin affinity chromatography of Nonidet P-40 (BDH Chemicals, Poole, England) cell extracts (Brown & Nathenson, 1977). Sephadex G-75 column chromatography as described by Ewenstein et al. (1978) was used to separate β_2m from H-2. In several cases, H-2K^d was isolated from glycoprotein pools from which H-2D^d (Nairn et al., 1980a) and/or H-2L^d had been previously removed (Coligan et al., 1980). The pools were treated with *Staphylococcus aureus* Cowan I strain cells after each immunoprecipitation to remove any residual immune complexes.

Preparation of CNBr Fragments. Cleavage with CNBr and isolation of peptides were carried out as described by Ewenstein et al. (1976) except that the CNBr fragments were isolated by chromatography over a Sephacryl S-200 column (2 × 200 cm) rather than a Sephadex G-100 column. Peptides were desalted as previously described by Ewenstein et al. (1978).

Digestion by Staphylococcus aureus V8 Protease. *Staphylococcal aureus* V8 protease (Miles Laboratories, Inc., Elkhart, IN) was used without further treatment. CNBr fragment E (containing 2 mg of cytochrome *c*) was dissolved in 0.9 mL of 0.5% NH_4HCO_3, pH 8.0, and 12 μL of V8 protease solution (1 mg/mL in 0.5% NH_4HCO_3, pH 8.0) was added. After incubation at 37 °C for 24 h, the reaction was halted by the addition of 0.1 mL of 88% formic acid, and the material was fractionated by a Sephadex G-25$_{sf}$ (superfine) column (0.9 × 220 cm) equilibrated in 2 M formic acid. In certain instances, an atypical chymotryptic-like cleavage was noted. Changing the V8 protease preparation from Miles Lot No. 0609 to Miles Lot No. 0877 eliminated this problem.

Thrombin Digestion. Thrombin (B grade, Lot No. 703053, Calbiochem, La Jolla, CA) was further purified as described by Lundblad (1971). Digestions were carried out according to a procedure described by Uehara et al. (1980). In brief, CNBr fragment D along with 1–2 mg of cytochrome *c* was dissolved in 1 mL of 0.1 M NH_4HCO_3, pH 8.5, and incubated at 37 °C for 20 h with 0.2–0.4 mL of thrombin solution [1.7 ΔOD_{247} units min^{-1} mL^{-1}, as described by Hummel (1959)]. The reaction was halted by the addition of 0.10 mL of 88% formic acid and applied directly onto a column of Sephadex G-50$_{sf}$ (0.9 × 220 cm) which had been equilibrated with 2 M formic acid.

Preparation of Tryptic Overlap Peptides. Trypsin digestion was carried out by using TPCK–trypsin (Worthington Biochemicals, Freehold, NJ). Isolated H-2K^d, labeled in [^{35}S]Met and [^{3}H]His, -Ile, -Thr, -Val, -Trp, and -Phe, was reduced and carboxamidomethylated (Brown et al., 1974). After desalting, the freeze-dried protein was dissolved in 0.1 M NH_4HCO_3, pH 8.5, at an unlabeled protein concentration of 10 mg/mL. Trypsin (10 mg/mL in 0.1 M HCl) was added to a final concentration of 3% and the mixture then incubated at 37 °C for 3.5 h. The reaction was halted by acidification with 88% formic acid, and the mixture was centrifuged to remove insoluble material and applied to a Sephadex G-25$_{sf}$ column (1.5 × 200 cm) equilibrated in 2 M formic acid. Pooled materials from the G-25$_{sf}$ eluate were freeze-dried, dissolved in 0.5 mL of 0.05% F_3AcOH, and further fractionated by high-performance liquid chromatography (HPLC). An E-M Lichrosorb (E. Merck, Darmstadt, West Germany) C-8 reverse-phase column (4.6 × 250 mm) was employed at a flow rate of 1 mL/min on a Waters Model 240 HPLC (Waters Associates, Inc., Milford, MA). Gradient elution from 0.05% F_3AcOH to 0.05% F_3AcOH/40% *n*-propyl alcohol was used to develop the chromatogram after an initial isocratic elution for 4 min. Radioactive peaks were pooled and freeze-dried, and the material was subjected to automated amino acid sequence analysis.

Automated Amino Acid Sequence Analysis. Details of the radiochemical amino acid sequencing methodology have been published elsewhere (Coligan et al., 1979; Uehara et al., 1980). Automated Edman degradations were performed with a Beckman 890C sequencer by using Beckman programs 102974 (DMAA), 121078 (0.1 M Quadrol), and 21980 (1 M Quadrol) in the presence of 1–2 mg of cytochrome *c*. The latter two programs were used after addition of a cold-trap accessory to the sequencer. Polybrene (Aldrich Chemical Co., Milwaukee, WI) was applied to the sequencer cup to prevent premature washout of smaller peptides (Klapper et al., 1978). Phenylthiohydantoin derivatives of radioactive amino acids were identified by cochromatography with unlabeled standard phenylthiohydantoin amino acid derivatives by HPLC (Hewlett-Packard, Model 1080B, Avondale, PA) using the isocratic elution conditions of Gates et al. (1979). The peaks were collected automatically by timed elution, and the radiolabeled phenylthiohydantoin amino acid derivative was identified by liquid scintillation counting of the freeze-dried samples in 5.0 mL of Biofluor (New England Nuclear, Boston, MA).

Results

Incorporation of Radioactivity into H-2K^d. The H-2K^d molecule was obtained from C14 tumor cells that had been cultured in the presence of either a single or a selected mixture of ^{3}H- or ^{35}S-labeled amino acids.

In certain cases, certain amino acids have been seen to be converted, in part, into other amino acids, thereby giving rise to a mixture of amino acid labels. The degree of incorporation of radioactivity into C14 cells and interconversion of certain amino acids by this cell line have been previously reported for the H-2D^d glycoprotein, isolated from the C14 cell line by Nairn et al. (1980a). In this study, we noted that, in addition to those interconversions previously reported, [^{3}H]Glu was converted to [^{3}H]Gln as well as [^{3}H]Pro and [^{3}H]Arg, yielding

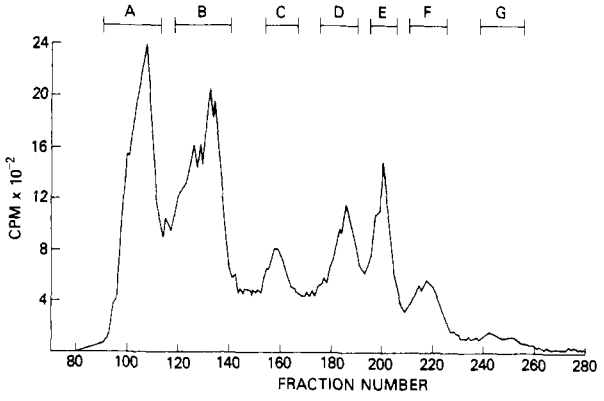

FIGURE 1: Sephacryl S-200 chromatography of [^{3}H]Ser-labeled CNBr peptides of H-2K^d. The column (2 × 200 cm) was equilibrated in 6 M Gdn·HCl. Fraction size was 2 mL at a flow rate of 0.2 mL/min.

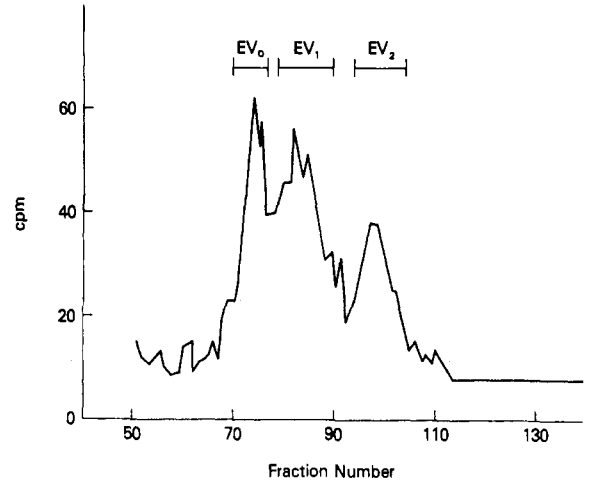

FIGURE 2: Sephadex G-25$_{sf}$ chromatography of V8 peptides of CNBr fragment E labeled in [^{3}H]Leu, -Tyr, and -Arg. The column (1 × 220 cm) was equilibrated in 2 M formic acid. Fraction size was 1 mL at a flow rate of 0.1 mL/min.

ratios of 1.0:0.3:0.2:0.16 Glu:Gln:Pro:Arg.

Isolation of CNBr Peptides. Sephacryl S-200 chromatography of the H-2K^d molecule after CNBr digestion yielded the profile shown in Figure 1. Some of these peptides had been identified previously (Kimball et al., 1980) after CNBr digestion of papain-treated H-2K^d and had been tentatively aligned by homology to the H-2K^b molecule. There was no papain treatment of immunoprecipitated H-2K^d in the present study. Pool A consists of partial cleavage products (Ewenstein et al., 1978; Kimball et al., 1980) whereas peptide pool B contains two disulfide-linked peptides. Peptides D and E constitute the amino-terminal 98 residues of the H-2K^d molecule (Kimball et al., 1980). Peptide C is a product of incomplete CNBr cleavage of the Met linking peptides D and E (Kimball et al., 1980). Peptides F and G had not been noted in the previous studies on the H-2K^d papain fragment, and partial amino acid sequence data indicate by homology to H-2K^b that they are derived from the carboxy-terminal portion of the H-2K^d molecule.

Enzymatic Cleavage of CNBr Fragment E. CNBr fragment E is the amino-terminal CNBr fragment of the H-2K^d molecule. Treatment of this peptide with staphylococcal V8 protease generated two major fragments, EV1 and EV2, as shown by gel filtration of the reaction mixture on a column of Sephadex G-25$_{sf}$ (Figure 2). Fragment EV0 was shown to be undigested material. The isolated V8 peptides were freeze-dried and then subjected to automated amino acid sequence analysis.

Table I: Absolute and Relative Amounts of Radioactivity Recovered in Fragment E V8 Peptides and Fragment D Thrombic Peptides Radiolabeled for Different Amino Acids

	V8 peptides									thrombic peptides								
		EV1		EV2		EV3				θ1		θ2		θ3			
labels[a,b]	initial cpm	radioactivity		radioactivity		radioactivity		% recovery[e]	initial cpm	radioactivity		radioactivity		radioactivity		% recovery[e]
		cpm[c]	%[d]	cpm	%	cpm	%			cpm[c]	%[d]	cpm	%	cpm	%	
His, Ile, Thr, Val, Trp, Phe	27 400	8 700	63.5	3000	21.9	2000	14.6	50	50 000	16 800	39.3	6400	15.0	19 500	45.7	85
Leu, Tyr, Arg	40 000	4 500	41.0	6400	59.0			27	30 000	6 000	50.0	2600	21.6	3 400	28.3	40
Glu, Gln, Pro, Arg	13 100	7 400	72.5	1800	17.6	1000	9.8	78	12 000	1 800	22.2	4500	55.6	1 800	22.2	68
Gln	4 700	<100	0	2000	100.0			43	20 000	6 800	35.2	8600	44.6	3 900	20.2	97
Ala	23 000	4 000	66.5	2000	33.5			26	15 000	3 900	100.0	<100	0	<100	0	26
Asn	12 000	<100	0	5100	100.0			43	18 000	8 500	100.0	<100	0	<100	0	47
Gly, Ser[f]	13 000	8 000	100.0	<100	0			62	13 000	2 000	66.6	1000	33.3	<100	0	23
Ser, Gly[g]	28 000	11 000	72.0	4300	28.0			55	28 000	8 600	54.7	2200	14.0	4 900	31.2	56
Pro	16 200	5 400	66.5	2700	33.5			50	ND	ND		ND		ND		
Lys	ND[h]	ND		ND		ND			16 200	4 000	69.0	<100	0	1 800	31.0	36

[a] No cysteine occurs in the above two CNBr peptides. [b] It has not been possible to biosynthetically incorporate [^{3}H]aspartic acid in vitro. [c] Radioactivity (cpm) recovered in each peptide. [d] Percent = (radioactivity in each peptide/total recovered radioactivity) × 100. [e] (Sum of cpm in all V8 or thrombic peptides/initial cpm) × 100. [f] Approximately 10% of [^{3}H]glycine is converted to [^{3}H]serine. [g] Approximately 30% of [^{3}H]serine is converted to [^{3}H]glycine. [h] ND, not determined.

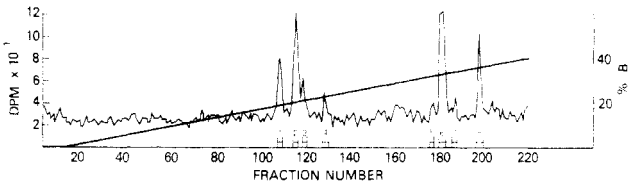

FIGURE 3: Reverse-phase high performance liquid chromatogram of tryptic pepties of CNBr fragment E labeled in [^{3}H]Ala, -Pro, -Glu, -Leu, -Tyr, and -Arg. A Lichrosorb C-8 column (4.6 × 250 mm) was used at a flow rate of 1 mL/min at 25 °c. Fractions of 0.4 mL were collected. The gradient was run over a period of 1 h, 20 min. Solvent A was 0.05% F_3AcOH. Solvent B was 0.05% F_3AcOH in *n*-propyl alcohol.

The recovery of radioactivity contained in each of the V8 peptides derived from the nine different labelings is summarized in Table I. Owing to the difficulty of separating undigested material, EV0, from peptide EV1, the relative yields of each peptide can only be approximated. However, the relative amounts of radioactivity in each peptide were, in general, in agreement with the relative amounts of each amino acid in the peptide. Furthermore, the failure to separate the undigested material (EV0) from peptide EV1 posed no problem for amino acid sequence determinations. Comparison of the amino acid sequence of EV1 to a partial sequence of intact fragment E extending 44 residues revealed its identity as the amino-terminal 30 residues. Therefore the undigested material (EV0) was combined with peptide EV1 for most amino acid sequence determinations.

Fragment EV2 was shown to comprise residues 31–52. On occasion, a chymotryptic-like cleavage at Phe-45 was noted, producing peptide EV3 which eluted after peptide EV2 and which spanned residues 46–52. Use of a different enzyme preparation (see Materials and Methods) eliminated the problem of unexpected cleavage.

Amino Acid Sequence of CNBr Fragment E. The complete amino acid sequence of fragment E is documented in Table I-S (see paragraph at the end of paper regarding supplementary material). The recovery of radioactivity contained in each step derived from automated amino acid sequence analyses is summarized in Figure 1-S (supplementary material). The carboxy-terminal residue of peptide EV1 is aspartic acid. Other investigators have also noted cleavage by this enzyme at certain aspartic acid residues (Drapeau, 1977; Gates et al., 1979). The absence of cleavage after the glutamic acid residues at positions 19 and 46 is probably due to the presence of proline residues at positions 20 and 47, respectively. Glu–Pro bonds have proven resistant to V8 protease cleavage in other proteins as well (Gates et al., 1979; Drapeau, 1977).

Peptide EV3 was determined to encompass positions 46–52 by virtue of the fact that it possessed a Trp at position 6 and the sequence Glu-Pro-Arg was noted in the first three residues of another preparation [multilabel group III, Tables I and I-S (supplementary material)] of EV3. Glu had previously been assigned to position 46 and Arg to position 48. Also, no Trp occurs in the first 45 residues of CNBr fragment E. Thus, peptide EV3 arose from a chymotryptic-like cleavage after Phe-45. The Trp assignment was confirmed by a tryptic overlap peptide, T6–14, which was derived from a preparation of the whole molecule labeled with [^{35}S]Met and [^{3}H]His, -Ile, -Thr, -Val, -Trp, and -Phe (see Figure 6).

In order to confirm certain assignments, fragment E, labeled with [^{3}H]Ala, -Pro, -Glu, -Leu, -Tyr, and -Arg, was digested by trypsin and subsequently fractionated into its components by reverse-phase HPLC (Figure 3). Material in the pools indicated was subjected to automated sequence analysis. Peptide ET2 was shown to span residues 45–48 and further

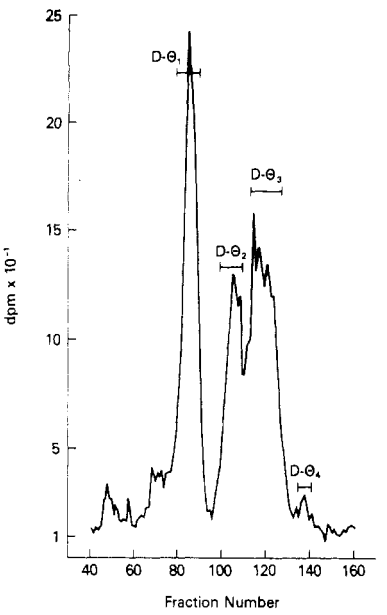

FIGURE 4: Sephadex G-50$_{sf}$ chromatography of thrombic peptides of CNBr fragment D labeled in [^{3}H]His, -Ile, -Thr, -Val, -Trp, and -Phe. The column (1 × 200 cm) was equilibrated in 2 M formic acid and run at a flow rate of 0.1 mL/min. One-milliliter fractions were collected and pooled as indicated.

confirms assignments in peptide EV3. Peptide ET4 was shown to encompass positions 49–52 by virtue of the fact that it was the only tryptic peptide labeled which began with Ala and which did not contain Arg. No Lys-containing tryptic peptides are possible since there is no Lys in CNBr fragment E. This Ala residue had previously been detected in peptide EV2 labeled only in Ala.

Positions 29, 30, 37, 39, and 41 have been assigned as aspartic acid because no other amino acid residues could be determined for those positions after labeling with all 19 other amino acids. The carboxy-terminal residue of fragment E and peptide EV2 is methionine based on a sequence analysis of tryptic peptide T6–14 (see Figures 5 and 6, below) derived from H-2K^d labeled with [^{35}S]Met and [^{3}H]His, -Ile, -Thr, -Val, -Trp, and -Phe and is consistent with the chemistry of CNBr cleavage (Kasper, 1975).

Enzymatic Cleavage of CNBr Fragment D. Fragment D, obtained after CNBr digestion of the H-2K^d molecule, is a glycopeptide spanning residues 53–98. Three major fragments (D-θ1, D-θ2, and D-θ3) were derived from this glycopeptide after treatment with the proteolytic enzyme thrombin as shown by the Sephadex G-50$_{sf}$ gel filtration profile (Figure 4). Often, peptide D-θ3 underwent a small amount of cleavage at the Arg at position 9, thereby generating an additional tetrapeptide, D-θ4 (Figure 4), and causing the elution profile of D-θ3 to be somewhat broadened. The recovery of radioactivity contained in each of the thrombic peptides from the nine different labelings is summarized in Table I and is in agreement with the relative amounts of each labeled amino acid found in the thrombic peptides. The isolated thrombic peptides were freeze-dried and then subjected to automated amino acid sequence analysis.

Comparison of the amino acid sequence of D-θ2 to a partial sequence of intact fragment D extending 40 residues revealed its identity as the amino-terminal 14 residues. Similar analyses of D-θ1 and D-θ3 allowed alignment of these two peptides and indicated that D-θ3 begins at residue 15 and that D-θ1 begins at residue 28. The identity of fragment D-θ1 as a glycopeptide was confirmed by finding that a preparation of fragment D labeled in [^{3}H]fucose yielded radioactivity only in D-θ1 upon

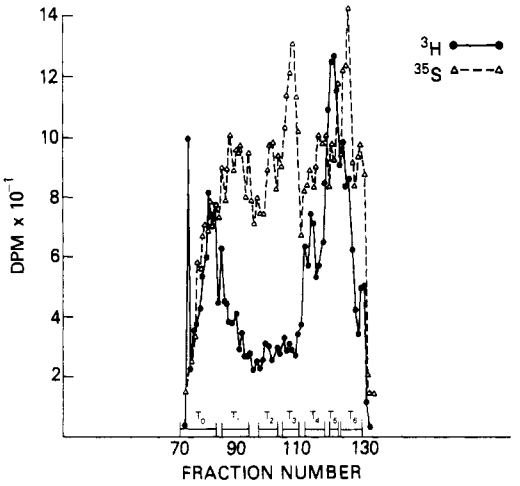

FIGURE 5: Sephadex G-25$_{sf}$ chromatography of tryptic overlap peptides of H-2K^d labeled in [^{35}S]Met and [^{3}H]His, -Ile, -Thr, -Val, -Trp, and -Phe. The column (1.5 × 220 cm) was equilibrated in 2 M formic acid and run at a flow rate of 0.1 mL/min. One-milliliter fractions were collected and pooled as indicated.

thrombin digestion (data not shown).

Amino Acid Sequence of CNBr Fragment D. The complete amino acid sequence of fragment D is documented in Table II-S (supplementary material). The radioactivity recovered in each step during sequence analyses is summarized in Figure 2-S (supplementary material). Position 34 is tentatively designated as Asn because no other amino acid was detected at this position, and peptide D-θ1 was the only product of thrombin treatment seen to label with Asn.

The most likely explanation for the failure to detect Asn at this position during sequence analysis relates to the likelihood of attachment of the carbohydrate moiety present in D-θ1 at this position. Such carbohydrate-linked residues are not extracted by butyl chloride during sequence analyses (Waterfield & Bridgen, 1975). Further suggestive evidence that residue 34 is the presumptive site of glycosylation comes from the knowledge that the amino acid sequence around position 34 would be Asn-Gln-Ser and that Asn-X-Ser/Thr serves as a recognition sequence for glycosylation (Marshall, 1972). Finally, the H-2K^b molecule (Coligan et al., 1979) and H-2D^b molecule (Maloy et al., 1981) apparently have a glycosylated Asn residue at the homologous position.

In agreement with the known proteolytic specificity of thrombin, Arg was detected as the carboxyl terminus of D-θ1, D-θ2, and D-θ3. Position 18 (residue 70) is presumed to be Asp because no other amino acid could be assigned to that position after labeling with the other 19 amino acids. The carboxyl terminus of CNBr fragment D is Met based on the chemistry of CNBr cleavage and the fact that no other residues were detected when sequence analyses were carried out past that point.

Alignment of CNBr Fragment E and CNBr Fragment D. H-2K^d, labeled with [^{35}S]Met and [^{3}H]His, -Ile, -Thr, -Val, -Trp, and -Phe, was reduced, carboxamidomethylated, and then treated with TPCK–trypsin; the resulting digest was partially purified over a column of G-25$_{sf}$ (Figure 5). Pool T6 provided the overlap peptide, T6–14, after further resolution by reverse-phase HPLC (Figure 6). This latter peptide had the following sequence: --Trp-Met-------Trp---Thr--.

Complete Amino Acid Sequence of Residues 1–98 of H-2K^d. The amino acid sequence of the amino-terminal 98 residues of the H-2K^d molecule, based on the foregoing data, is shown in Figure 7.

Discussion

The complete amino acid sequence of the amino-terminal 98 residues of the H-2K^d major histocompatibility complex antigen is presented here. CNBr digestion provided two peptides which comprise the amino-terminal 98 residues. Fragment E comprises the first 52 residues in the amino terminus whereas fragment D extends from residue 53 to residue 98 and contains one of the glycosyl moieties of the H-2K^d molecule. Automated radiosequence analysis, applied in accordance with criteria outlined by Uehara et al. (1980), of staphylococcal V8 protease and thrombic and tryptic peptides of fragments E and D allowed assignment of the amino acid sequence given in Figure 7.

The true biological function of the H-2 alloantigens has not yet been unequivocably established. However, present evidence

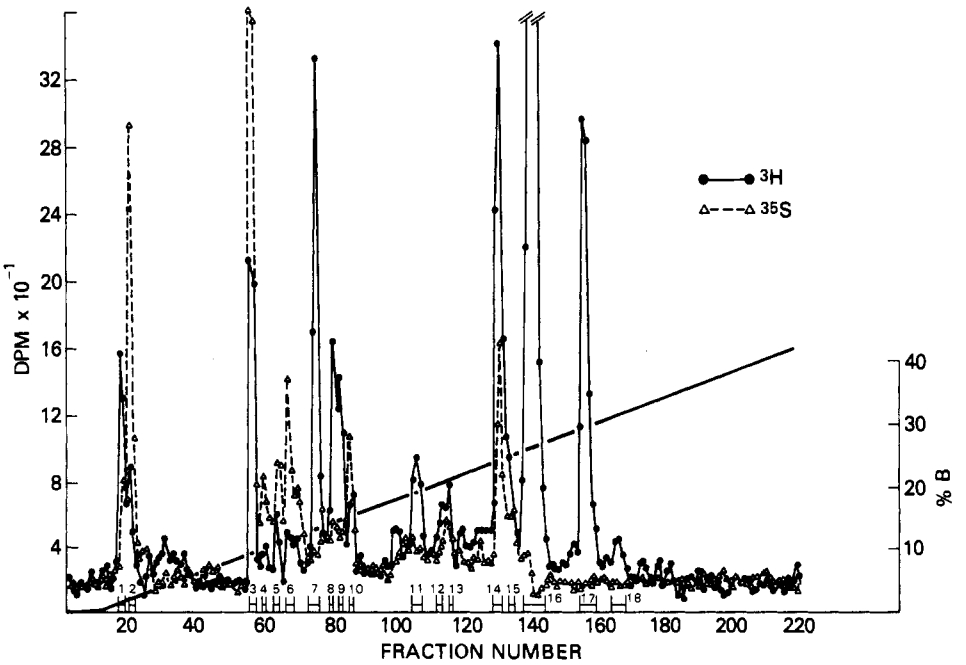

FIGURE 6: Reverse-phase high-performance liquid chromatogram of tryptic overlap peptide pool T$_6$ labeled in [^{35}S]Met and [^{3}H]His, -Ile, -Thr, -Val, -Trp, and -Phe. The same column conditions described in Figure 3 were used.

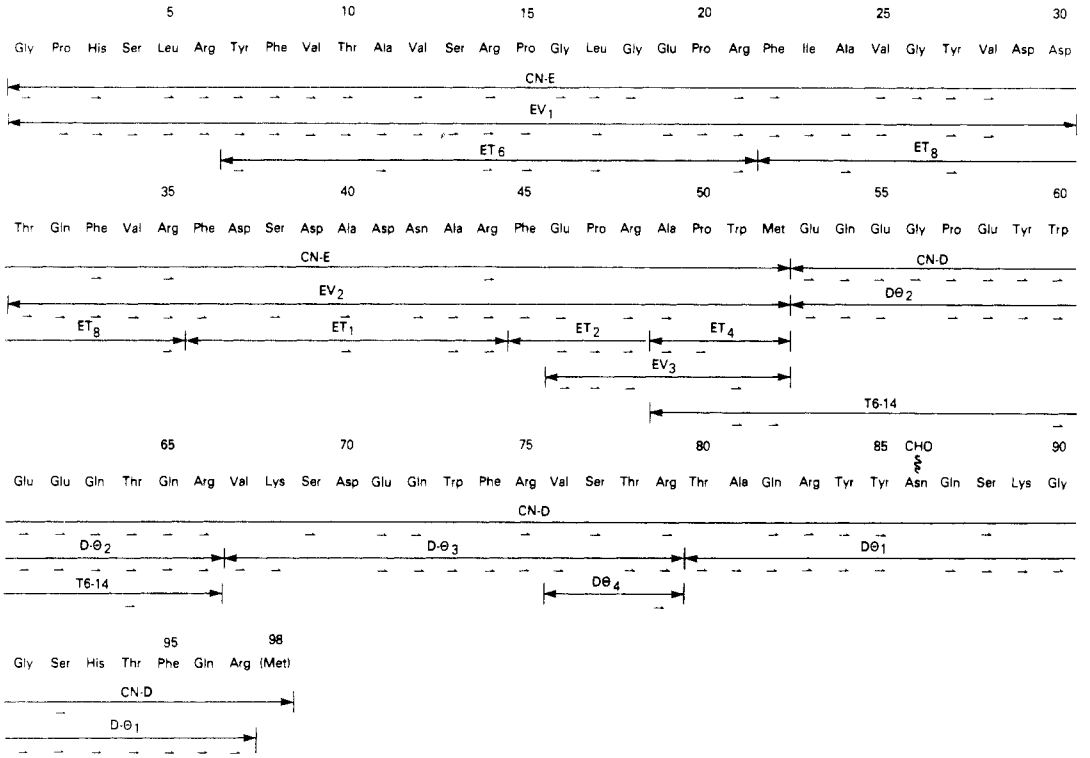

FIGURE 7: Amino acid sequence of residues 1–98 of the H-2K^d molecule. CN (CNBr), D-θ (thrombic), EV (staphylococcal V8 protease), ET (tryptic), and T (tryptic overlap) peptides used to obtain the sequence. Arrows below the line indicate which amino acids were identified in each peptide. Residues in italics were assigned indirectly. CHO indicates the position of the glycosyl unit.

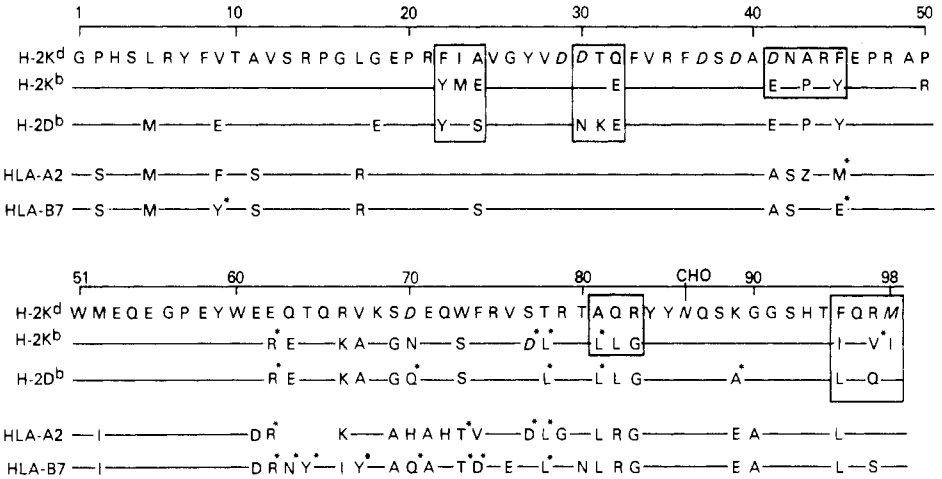

FIGURE 8: Comparison of the NH_2-terminal 98 residues of the H-2K^d, H-2K^b (Uehara et al., 1980), H-2D^b (Maloy et al., 1980), HLA-A2, and HLA-B7 (Orr et al., 1979) molecules. A solid line indicates identity to the H-2K^d sequence. An asterisk indicates amino acid interchanges which require more than a single base change. Boxes indicate clusters of amino acid differences, other than those in the region between residues 62 and 83, where the murine molecules differ from one another and which might also contribute to alloreactivity.

suggests that these products play an important role in T cell recognition of the cell membrane antigens through involvement in a process termed "associative recognition" (Paul & Benaceraff, 1977; Zinkernagel & Doherty, 1979). Their remarkable polymorphism as seen in serological analyses (Snell et al., 1976), peptide map comparisons (reviewed by Vitetta & Capra, 1978), and now NH_2-terminal sequence comparisons (Figure 8) is thought to reflect the role of specific recognition in this process. Defining the sites and nature of the amino acid differences between H-2 allelic products could be expected to provide information concerning the nature of these recognition processes and, in addition, to indicate regions of these molecules that might bear the epitopes recognized by alloantisera. The comparison of the amino acid sequence of the first 98 residues of H-2K^d to the homologous regions of H-2K^b, H-2D^b, HLA-A2, and HLA-B7 reveals an interesting arrangement both of differences and similarities. As shown in Figure 8 and tabulated in Table II, H-2K^d shows an overall homology of 77% to H-2K^b and 75%, 71%, and 68% homology to H-2D^b, HLA-A2, and HLA-B7, respectively. Of interest is the finding that the majority of the amino acid substitutions which occur between H-2K^d and H-2K^b are located between residues 62 and 83 (45% homology). A similar finding is apparent when H-2K^d is compared with H-2D^b, HLA-A2, and HLA-B7, which show 50%, 25%, and 26% homology, respectively. In addition, 5 of the 15 amino acid substitutions which occur between H-2K^d and H-2K^b, between residues 62 and 98, require multiple base changes at the nucleic acid level whereas all of the remaining substitutions in the amino-terminal 98 residues can be attributed to single nucleic acid base changes. It is tempting to speculate, therefore, as has been pointed out previously in comparative studies in the HLA series

Table II: Percent Homology of H-2K^d to Other H-2 and HLA Antigens between Residues 1 and 98[a]

	H-2K^d		H-2K^b		H-2D^b		HLA-A2	
	1–98	62–83	1–98	62–83	1–98	62–83	1–98	62–83
H-2K^b	77(5)[b]	45 (4)						
H-2D^b	75(5)	50(4)	84	90				
HLA-A2	71	25[c]	71	50[c]	68	47[c]		
HLA-B7	68	26[c]	66	39[c]	72	52[c]	86	40[d]

[a] H-2K^d data, this paper; H-2K^b from Uehara et al. (1980); H-2D^b from Maloy et al. (1981); HLA-A2 and HLA-B7 from Orr et al. (1979). [b] Number in parentheses indicates the number of multiple base changes required to account for some of the amino acid differences. [c] Residues 61–83. [d] Residues 66–80.

(Orr et al., 1979), that the dramatic amino acid sequence differences in the region between residues 62 and 83 reflect molecular differences responsible for the specific biological recognition function of the H-2 antigens.

Although positions 62–83 are the most variable regions when H-2 alloantigens are compared, it is important to note that other parts of the molecule also show discrete regions of difference. For example, inspection of Figure 8 reveals that H-2K^d contains five small clusters outlined in boxes which have unique amino acid sequences. The clusters located between residues 22 and 24, 30 and 32, and 95 and 98 show unique sequences for all three molecules. The clusters which include positions 41–45 and 81–83 are unique to H-2K^d. It may also be noteworthy that both H-2K^d and H-2K^b have the sequence Asp-Asp at residues 29–30 whereas H-2D^b and H-2D^d (R. N. Nairn, S. G. Nathenson, and J. E. Coligan, unpublished results) have the sequence Asp-Asn instead. Whether these constitute K-end or D-end epitopes remains to be seen. It seems likely, although certainly not proven, that these clusters of amino acid differences are involved in the allodeterminants of these molecules. Whether these clusters of amino acid differences actually constitute multiple allodeterminants or whether some or all of them combine into three-dimensional arrays to comprise larger epitopes remains to be determined.

Whereas the differences between alleles is quite striking, it seems apparent from recent data on structural characterization of H-2 mutants that only small changes are needed for immune recognition (Ewenstein et al., 1980). For example, studies have already suggested that in the simplest case, differences between parental and mutant products as small as one or two amino acid interchanges, often widely separated, are sufficient to elicit a highly responsive and specific immune recognition, both in alloaggression and in associative recognition (reviewed by Nairn et al., 1980b). Therefore, the extraordinary differences between the allelic products noted above provide a potentially enormous repertoire of recognitive determinants which can provide the basis of T cell allorecognition and serological distinctions.

Previously, peptide-mapping studies [see review by Vitetta & Capra (1978)] had indicated that the extent of structural differences and similarities between H-2K or D allelic products was as great as those that exist between products of different loci; i.e., there was no evidence of "K-ness" or "D-ness". The present data obtained for H-2K^d permit the first extensive direct comparison of primary structures of allelic histocompatibility antigen gene products. The data support those early conclusions obtained by peptide mapping. Comparisons of the primary structure of the first 98 residues of the H-2K^d and H-2K^b alleles and H-2D^b gene products reveal no overall discernible structural features which would readily allow one to differentiate a K-end gene product from a D-end gene product. In fact, as shown in Figure 8, H-2K^b more closely resembles H-2D^b than H-2K^d. Also, studies on H-2D^d (R. Nairn, S. G. Nathenson, and J. E. Coligan, unpublished data) indicate that it is no more homologous to H-2D^b than to H-2K^d or H-2K^b.

While our present data only cover an analysis of the first 98 residues, these findings suggest that the allelic relationships and the geneological understanding of these gene products and their controlling genes will not be simple. Clearly, gene cloning techniques which will permit direct analysis of the genes themselves are required to more fully expolore the interrelationships.

Added in Proof

Recently we have developed the capability to biosynthetically incorporate [^{3}H]Asp (F. T. Gates, III, E. S. Kimball, and J. E. Coligan, unpublished results). We have thus far confirmed the assignments of aspartic acid at residues 29, 30, 37, 39, and 41.

Acknowledgments

We express our gratitude to Pat Spinella and Robert Valas for excellent technical assistance and to Virginia Frye for excellent secretarial assistance. Discussions with Drs. T. J. Kindt, W. L. Maloy, J. M. Martinko, R. Nairn, and H. Uehara were most helpful and are gratefully acknowledged.

Supplementary Material Available

Two tables (I-S and II-S) showing the identification of the amino acids in CNBr fragments E and D and two figures (1-S and 2-S) showing recovery of radioactivity in each step of automated sequence analyses of the peptides comprising the first 98 residues of H-2K^d (8 pages). Ordering information is given on any current masthead page.

References

Brown, J. L., & Nathenson, S. G. (1977) *J. Immunol. 118*, 98.

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.

Coligan, J. E., Kindt, T. J., Nairn, R., Nathenson, S. G., Sachs, D. H., & Hansen, T. H. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1134.

Cresswell, P., Turner, M. J., & Strominger, J. L. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 1603.

Drapeau, G. R. (1977) *Methods Enzymol. 47*, 189.

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.

Ewenstein, B. M., Uehara, H., Nisizawa, T., Melvold, R. W., Kohn, H. I., & Nathenson, S. G. (1980) *J. Immunol. 11*, 383.

Gates, F. T., III, Coligan, J. E., & Kindt, T. J. (1979) *Biochemistry 18*, 2267.
Hummel, B. C. W. (1959) *Can. J. Biochem. Physiol. 37*, 1393.
Kasper, C. B. (1975) *Mol. Biol. Biochem. Biophys. 8*, 115.
Kimball, E. S., Maloy, W. L., Martinko, J. M., Nathenson, S. G., & Coligan, J. E. (1980) *Mol. Immunol. 17*, 1283.
Klapper, D., Wilde, C. W., & Capra, J. D. (1978) *Anal. Biochem. 85*, 126.
Klein, J. (1975) *Biology of the Mouse Histocompatibility Complex*, Springer-Verlag, New York.
Lundblad, R. L. (1971) *Biochemistry 10*, 2501.
Maloy, W. L., Nathenson, S. G., & Coligan, J. E. (1981) *J. Biol. Chem. 256*, 2863.
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 19*, 6188.
Nairn, R., Nathenson, S. G., & Coligan, J. E. (1980a) *Eur. J. Immunol. 10*, 495.
Nairn, R., Yamaga, K., & Nathenson, S. G. (1980b) *Annu. Rev. Genet. 14*, 241.
Nakamuro, K., Tanigaki, N., & Pressman, D. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 2863.
Orr, H. T., Lopez de Castro, J. A., Parham, P., Ploegh, H. L., & Strominger, J. L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4395.
Paul, W. E., & Benaceraff, 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.
Shimada, A., & Nathenson, S. G. (1969) *Biochemistry 8*, 4048.
Snell, G. D. (1968) *Catalog of Mouse Alloantisera*, Transplantation Immunology Branch, National Institutes of Health, Bethesda, MD.
Snell, G. D., Dausett, J., & Nathenson, S. G. (1976) *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.
Waterfield, M. D., & Bridgen, J. (1975) in *Instrumentation in Amino Acid Sequence Analysis* (Perham, R. N., Ed.) p 41, Academic Press, New York.
Zinkernagel, R. M., & Doherty, P. C. (1979) *Adv. Immunol. 27*, 51.

Effect of the Proton Electrochemical Gradient on Maleimide Inactivation of Active Transport in *Escherichia coli* Membrane Vesicles†

Dunell E. Cohn, Gregory J. Kaczorowski,‡ and H. Ronald Kaback*

ABSTRACT: Many active transport systems present in *Escherichia coli* membrane vesicles are inhibited by maleimides and other sulfhydryl reagents. These reagents do not interfere with the oxidation of reduced phenazine methosulfate or with the electrochemical proton gradient ($\Delta\bar{\mu}_{H^+}$). The rate of inactivation of the β-galactoside transport system by various maleimides is increased in the presence of reduced phenazine methosulfate, and it is shown that the electrochemical proton gradient is responsible for the effect. Furthermore, similar effects are observed with the proline and melibiose transport systems. Thus, it appears that either the reactivity or accessibility of a sulfhydryl group(s) in each of these carriers is altered by the presence of a transmembrane $\Delta\bar{\mu}_{H^+}$. The findings are consistent with the notion that $\Delta\bar{\mu}_{H^+}$, in addition to acting as the immediate driving force for active transport, may bring about structural or conformational changes in certain membrane proteins that catalyze active transport.

The chemiosmotic hypothesis of Mitchell (1961, 1966, 1968, 1973, 1979) proposes that energy derived from respiration is transformed into a transmembrane electrochemical proton gradient ($\Delta\bar{\mu}_{H^+}$)[1] that is the immediate driving force for many active transport systems in bacterial cells. Accordingly, respiration-dependent accumulation of substrates is postulated to result from the obligatory coupling of substrate translocation with proton translocation mediated by substrate-specific membrane proteins (carriers or porters).

Cytoplasmic membrane vesicles prepared from *Escherichia coli* by osmotic lysis (Kaback, 1971; Short et al., 1975) have the same polarity of the membrane as the intact cell (Owen & Kaback, 1978, 1979a,b) and retain the capacity to catalyze the active transport or many substrates (Kaback, 1974a; Ramos et al., 1976). Studies performed both with intact cells (Harold, 1976) and with membrane vesicles (Ramos & Kaback, 1977a-c; Konings & Boonstra, 1977) have provided virtually unequivocal evidence for the central role of chemiosmotic phenomena in respiration-dependent active transport.

Kinetic studies of lactose transport in *E. coli* membrane vesicles have provided some interesting insights into the effects

† From the Laboratory of Membrane Biochemistry, Roche Institute of Molecular Biology, Nutley, New Jersey 07110. *Received October 15, 1980.*

‡ Fellow of the Helen Hay Whitney Foundation. Present address: Department of Biochemistry, Merck Sharp & Dohme, Rahway, NJ 07065.

[1] Abbreviations used: DTT, dithiothreitol; PMS, phenazine methosulfate; TMG, methyl 1-thio-β-D-galactopyranoside; NNM, *N*-(2-naphthyl)maleimide; MBTA, (4-(*N*-maleimido)benzyl)trimethylammonium iodide; GSMal, glutathione maleimide; NEM, *N*-ethylmaleimide; pBM, *N*-(*p*-benzoyl)maleimide; TSM, *N*-(*p*-tolylsulfonyl)-maleimide; $\Delta\bar{\mu}_{H^+}$, electrochemical gradient of protons; ΔpH, pH gradient; $\Delta\Psi$, membrane potential.

0006-2960/81/0420-3308$01.25/0