A Nonpolymorphic Major Histocompatibility Complex Class Ib Molecule Binds a Large Array of Diverse Self-peptides

By Sebastian Joyce,*¶ Piotr Tabaczewski,‖ Ruth Hogue Angeletti,§ Stanley G. Nathenson,*‡ and Iwona Stroynowski‖

*From the Departments of *Cell Biology and ‡Microbiology and Immunology, and the §Laboratory of Macromolecular Analysis and the Department of Developmental Biology and Cancer, Albert Einstein College of Medicine, Bronx, New York 10461-1975; the ‖Department of Microbiology and the Center for Diabetes Research, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75235-8854; and the ¶Department of Microbiology and Immunology, Milton S. Hershey Medical Center, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania 17033*

Summary

Unlike the highly polymorphic major histocompatibility complex (MHC) class Ia molecules, which present a wide variety of peptides to T cells, it is generally assumed that the nonpolymorphic MHC class Ib molecules may have evolved to function as highly specialized receptors for the presentation of structurally unique peptides. However, a thorough biochemical analysis of one class Ib molecule, the soluble isoform of Qa-2 antigen (H-2SQ7^b), has revealed that it binds a diverse array of structurally similar peptides derived from intracellular proteins in much the same manner as the classical antigen-presenting molecules. Specifically, we find that SQ7^b molecules are heterodimers of heavy and light chains complexed with nonameric peptides in a 1:1:1 ratio. These peptides contain a conserved hydrophobic residue at the COOH terminus and a combination of one or more conserved residue(s) at P7 (histidine), P2 (glutamine/leucine), and/or P3 (leucine/asparagine) as anchors for binding SQ7^b. 2 of 18 sequenced peptides matched cytosolic proteins (cofilin and L19 ribosomal protein), suggesting an intracellular source of the SQ7^b ligands. Minimal estimates of the peptide repertoire revealed that at least 200 different naturally processed self-peptides can bind SQ7^b molecules. Since Qa-2 molecules associate with a diverse array of peptides, we suggest that they function as effective presenting molecules of endogenously synthesized proteins like the class Ia molecules.

The MHC class I molecules serve as receptors for the transport and display of endogenously derived self- and non-self-peptides at the cell surface. Non-self-peptides (e.g., derived from intracellular pathogens), when presented by class I molecules, form ligands for TCRs triggering antigen-specific lysis of the infected cells by CTL (reviewed in reference 1). The presentation of self- and non-self-peptides is attributed to the ubiquitously expressed, highly diverse classical transplantation antigens known as the class Ia molecules, which in the mouse are encoded by the *H-2K* and *D* region genes. The diversity of the class Ia molecules mainly affects the antigen binding groove, which is thought to be important for presenting a wide variety of peptides for immune surveillance by T cells (2). It is intriguing, however, that the majority of the class I genes in the mouse MHC (>30) distributed over the *H-2Q*, *T*, and *M* regions are not polymorphic. Their products, known as the nonclassical class Ib molecules, compared with the class Ia molecules are highly conserved, are expressed in a tissue specific manner, and their physiological role(s) is unknown (3, 4).

Several recent studies have shown that H-2Q, T, and M molecules can serve as weak transplantation antigens in vivo and in vitro (3–6). Further, they can also present intracellular pathogen-derived antigens to specific CTL (7–11). In several instances, the CTL responses are peptide dependent (7–9, 12–16), suggesting that one of the physiological functions of class Ib molecules might be to control immune responses to intracellular pathogens akin to the class Ia molecules. Since, the antigen binding grooves of the class Ib molecules are highly conserved, it is often thought that they may have evolved to present a limited set of unique peptides. This has been shown to be true for H-2M3 molecule, which requires an N-formylated peptide for binding, a feature present only in 13 mitochondrial (self)-proteins (7, 8, 12–16). Whether this is a general feature for class Ib binding peptides is not known.

J. Exp. Med. © The Rockefeller University Press • 0022-1007/94/02/0579/10 $2.00
Volume 179 February 1994 579–588

Qa-2 antigens are among the best characterized class Ib molecules. Encoded by two almost identical genes, *Q7* and *Q9* in the *b* haplotype mice, they are expressed on lymphoid cells and primitive hematopoietic progenitors in the bone marrow in adult mice (17). They are noncovalently associated with β_2-microglobulin (β_2-m)[1] and exist in vivo in two forms: the 40-kD Qa-2 antigen anchored to the membrane by glycophosphatidylinositol (GPI) (8, 19) and the 39-kD soluble molecule (20), SQa-2. The switch from membrane-bound to soluble form is induced upon activation of the immune system and is thought to play a role in the regulation of the immune response (20, 20a).

In an approach to solve the major puzzle as to the role of the numerous nonpolymorphic MHC class Ib molecules, we have analyzed the biochemical features of the soluble isoform of Qa-2 antigen, $SQ7^b$. We find that $SQ7^b$ molecules consist of heavy and light chains complexed with nonameric peptides in 1:1:1: ratio. These peptides contain a $Q7^b$-specific binding motif and are derived from intracellular proteins. These data and the minimal estimates of $SQ7^b$ binding peptide repertoire support the conclusion that Qa-2 molecules bind a diverse array of naturally processed peptides in much the same manner as the highly polymorphic MHC class Ia molecules.

Materials and Methods

High Expression Cell Line. Large quantities of SQa-2 complexes were produced in NS0 myeloma cells using pHEKmDHFR vector-based expression system (21). The cDNA encoding $SQ7^b$ (20) was cloned into the SalI site of pHEKmDHFR downstream from the immunoglobulin heavy chain enhancer and κ chain promoter splice site cassette (Fig. 1) to ensure efficient expression of the recombinant cDNA. The cDNA differs at two positions from the reported sequence (22; and Goyarts, E. C., and S. G. Nathenson, unpublished results); their location does not affect the peptide binding groove nor β_2-m interaction. Downstream from the cDNA is the 3′ end of the κ constant region and its polyadenylation site. The vector also contains a DNA amplification selection marker, mutant dihydrofolate reductase, which binds methotrexate (Mtx) poorly such that the transfected cells can be selected in concentrations of Mtx that would be normally lethal to the cells. Genes expressed in this manner in the pHEKmDHFR vector have been shown to be coamplified, resulting in high expression of the recombinant protein in myeloma cells (21). The recombinant pHEKmDHFR-sQ7b plasmid was introduced into NS0 cells by electroporation. Mtx-resistant transfectants were selected in DMEM supplemented with 5% FCS and 0.2 μM Mtx. The transfected pool was subjected to increasing concentrations of Mtx over a period of few months essentially as described by Hendricks et al. (21). Cells resistant to 10 μM Mtx were cloned and $SQ7^b$ high expressors were identified by ELISA. Selected clones were then cultured with increasing concentrations of Mtx. The highest $SQ7^b$ secreting clone resistant to 64 μM of Mtx was grown in a CelliGen™-stirred tank bioreactor (New Brunswick Sci., Edison, NJ) in 1:1 DMEM/RPMI supplemented with 3% FCS, 1 mM sodium pyruvate, 0.01% pluronic F.66, and 0.6 μM Mtx. Under these conditions, the cells grown with agitation of 40 rpm produced up to 10 mg of $SQ7^b$/10 liters of tissue culture–spent medium over a period of 2–5 d. The spent medium thus obtained was used for $SQ7^b$ and peptide analysis.

ELISA. ELISA was performed by a modification of the method described by Harlow and Lane (23). Briefly, polyvinyl chloride 96-well plates were coated with Qa-2-specific affinity-purified mAb M46 (24). The uncoupled sites were blocked with 3% BSA diluted in PBS, pH 7.4. Tissue culture supernatants containing $SQ7^b$ or purified $SQ7^b$ in PBS were added to the wells and incubated overnight at 4°C. Supernatants from the wells were aspirated and the wells were washed three times with 0.5% Tween-20 in PBS (PBS/Tween). Biotinylated Qa-2-specific mAb 20-8-4s (25) was added to the $SQ7^b$-coated wells and incubated for 4 h at 4°C. The plates were washed six times with PBS/Tween. The reaction was amplified with β-galactosidase-conjugated streptavidin (Boehringer Mannheim Biochemicals, Indianapolis, IN) for 30 min at 4°C, washed six times as above, and developed using the β-galactosidase substrate red-β-D-galactopyranoside (Boehringer Mannheim Biochemicals). The concentration of $SQ7^b$ in the supernatants was estimated by recording absorbance of the red product at 570 nm using a microtiter plate reader (Bio-Tek Instrs., Winooski, VT).

Affinity Chromatography. $SQ7^b$ molecules were purified from culture supernate of NS0 transfectants. The supernate was first concentrated to 1:10 the volume by saturated ammonium sulfate precipitation (final concentration of 80%). The concentrate was quantitated using conformation-dependent mAb 20-8-4s and M46 in an ELISA (data not shown). The $SQ7^b$ was affinity purified using 3.2 ml 20-8-4s-coupled protein A–Sepharose column. Affinity columns were prepared by coupling the protein A (Pharmacia LKB, Gaithersburg, MD)–purified antibody to protein A–Sepharose using dimethyl pimelimidate-2 HCl (Pierce, Rockford, IL) as described by Harlow and Lane (23). The concentrate (120 ml) was passed over normal mouse serum, B22-249 (D^b specific; 26), and 20-8-4s columns, in that order, once at the rate of 0.5 ml/min. The B22-249 and 20-8-4s columns were washed with 20-column volumes of PBS and eluted with 5 ml of 0.1 N acetic acid (pH 2.9); the first 1 ml was discarded and the next 4.5 ml was collected. The eluted materials were denatured by increasing the concentration of acetic acid to 2.0 N (pH 1.96) followed by incubation of the mixture in a boiling water bath for 5 min. The mixture was cooled for 15 min at room temperature. A 100-μl aliquot was used for amino acid analysis and another 100 μl was subjected to amino acid sequence analysis. Data from both these analyses were used to quantitate the initial amount of $SQ7^b$ molecules. Peptides from 4.0 ml of the denatured material were separated by centrifugation using a Centricon 3 filtration unit (Amicon Corp., Danvers, MA). The filtrate so obtained is the acid-eluted fraction. The acid-eluted fraction was concentrated to 120 μl in a speed-vac (Savant Instrs., Inc., Farmingdale, NY), of which 10% was used to quantitate peptide yield by amino acid analysis.

Reversed-Phase (RP) HPLC. The remainder of the acid-eluted ultrafiltrate (90%; ~2.7 nmol) was loaded onto a 1.0 × 250-mm nucleosil C18 column (Alltech Assoc. Inc., Deerfield, IL) and separated by RP-HPLC using HP1090 (Hewlett-Packard Co., Palo Alto, CA) equipped with an on-line 1040 diode array detector. The column was eluted using 0.06% trifluoroacetic acid, 5% acetonitrile, water (buffer A) and 0.05% trifluoroacetic acid, 80% acetonitrile, water (buffer B). The gradient of the mobile phase consisted of 0% B at start (0.01 min), 37% B at 63 min, 70% B at 95 min, 90% B at 105 min, and 100% B at 110 min, established at a rate of 75 μl/min. Chromatographic data was acquired at 210 and 280 nm. Eluting peaks were monitored at 210 nm and individual peaks

[1] *Abbreviations used in this paper:* β_2-m, β_2-microglobulin; Mtx, methotrexate; RP, reversed phase.

were manually collected after a lag time of 0.2 min. Each fraction was quick frozen on dry ice and stored at −70°C until further use. Repurification of individual peaks was performed essentially as above. The elution gradient consisted of either of two conditions: peaks 3, 5, 7, 10, and 19 were purified using 0% B at start (0.01 min), 37% B at 63 min, and 100% B at 70 min, at a flow rate of 50 μl/min; the remainder of the peaks were repurified using 20% B at 10 min, 45% B at 70 min, and 100% B at 75 min, at a flow rate of 50 μl/min. Peptide preparation from B22-249 column–eluted material was separated by RP-HPLC using the first gradient. The chromatogram so obtained was used to determine the specific peaks obtained from peptides eluted from $SQ7^b$.

Amino Acid Analysis and Sequencing. The peptide or protein preparations were acid hydrolyzed in 6 N HCl in vacuo at 110°C for 24 h. The hydrolysate was speed-vac concentrated and made to 6 μl in distilled water. A fraction (1 μl) was analyzed using HP AminoQuant (Hewlett-Packard Co.) system with *o*-phthalaldehyde and f-moc precolumn derivatization. Fluorescence data was collected using an HP 1046 detector (Hewlett-Packard Co.). Quantitation of the affinity-purified $SQ7^b$ complex using the yields of alanine (n = 27.5), arginine (n = 25.5), histidine (n = 14), and leucine (n = 31.5) from the amino acid analysis data resulted in 5.75 nmol of the starting material from 120 ml of crude concentrate. A similar analysis of the acid-eluted ultrafiltrate yielded 3.03 nmol of peptides using the yields of histidine and leucine as the common invariant residues in the peptide after correcting for nonspecific background. The background for this analysis was determined as the amount of histidine and leucine present in the ultrafiltrate of the material eluted from B22-249 (isotype-matched nonspecific mAb) column.

Amino acid sequencing was performed by automated Edman degradation (ABI 477A; Applied Biosystems, Inc., Foster City, CA). The arginine residues in the heavy chain of $SQ7^b$ were ascertained by sequence analysis using a Porton 2000 sequencer (Beckman Instrs., Fullerton, CA).

A data base search was performed to compare the peptide sequences obtained with those stored in the Swiss Protein Data Bank (SwissProt release 25.0; 4/93) using the Wisconsin Genetics Computer Group (GCG) sequence analysis package (27).

Results

Overexpression of SQa-2 Molecules in Mammalian Cells. Since Qa-2 molecules are expressed in vivo as a heterogeneous mixture of products of several related genes in relatively low quantities (28), we have developed a genetically engineered cell line producing high levels of a single Qa-2 product, $SQ7^b$. Full-size cDNA encoding $SQ7^b$ (20) was subcloned into the pHEKmDHFR vector (21), introduced into NS0 cells by electroporation, and selected by stepwise methotrexate treatment (Fig. 1). NS0 cells (GPI- and SQa-2-negative BALB/c-derived myeloma) were used so that the host cell producing the recombinant protein will be closely related to the cells expressing Qa-2 in vivo, which is important in the analysis of endogenous self-peptides. The highest SQa-2-secreting clones were estimated to produce ~0.5–1.0 mg of $SQ7^b$ per liter of spent tissue culture medium when grown for 2–3 d in a bioreactor.

The SQa-2 Peptide Binding Motif Is Unique. The secreted form of $Q7^b$ was purified from spent tissue culture medium. The spent medium containing $SQ7^b$ was concentrated 10-fold by saturated ammonium sulfate precipitation. The integrity of the $SQ7^b$ molecules in the concentrate when tested by ELISA using a conformation-dependent mAb showed no loss of $SQ7^b$ structure. $SQ7^b$ molecules were further purified by affinity chromatography using 20-8-4s mAb and the resulting Qa-2 proteins were tested by SDS-PAGE (data not shown) and/or amino acid sequence analysis (see below). Peptides constitutively bound to $SQ7^b$ were then purified from the affinity column eluate by acid denaturation followed by the separation of the low molecular weight peptides (acid-

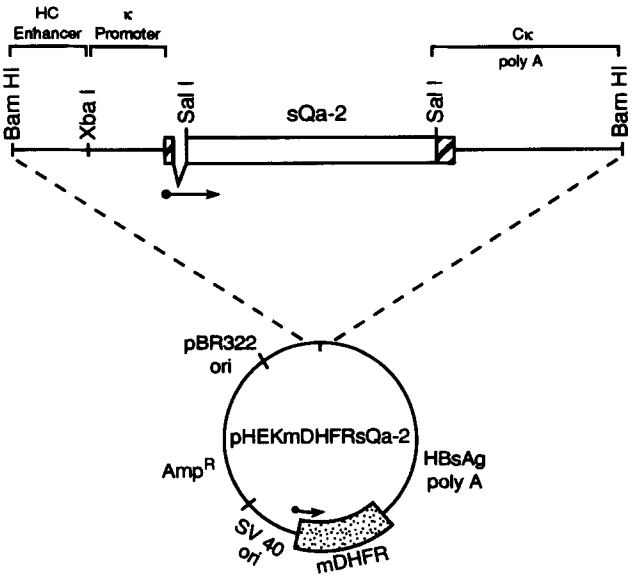

Figure 1. High expression vector containing $SQ7^b$ cDNA used for overexpression of SQa-2 molecules in a mammalian myeloma cell line, NS0. The designations in pHEKmDHFR are from Hendricks et al. (21), and of SQa-2 are from Ulker et al. (20).

Table 1. *$SQ7^b$ Binding Motif of Constitutively Bound Peptides*

	Position of residue								
	P1	P2	P3	P4	P5	P6	P7	P8	P9
Predominant residue*							H		
	K‡	Q	L	P	I	K		E	L
	A	L	N	E	V	L		Q	I
	G			G	Q	F		N	F
				K	E	M		K	
					L	Y		S	
					T			T	
					Y			V	

* Predominant residue is defined as the only residue that had a significant PTH amino acid yield at a given sequencing cycle.

‡ All sequencing cycles, except cycle 7, showed significant signals for two or more PTH amino acids, and the residue assignments for each cycle (position in the peptide) are listed in the order of decreasing yields of the PTH amino acid.

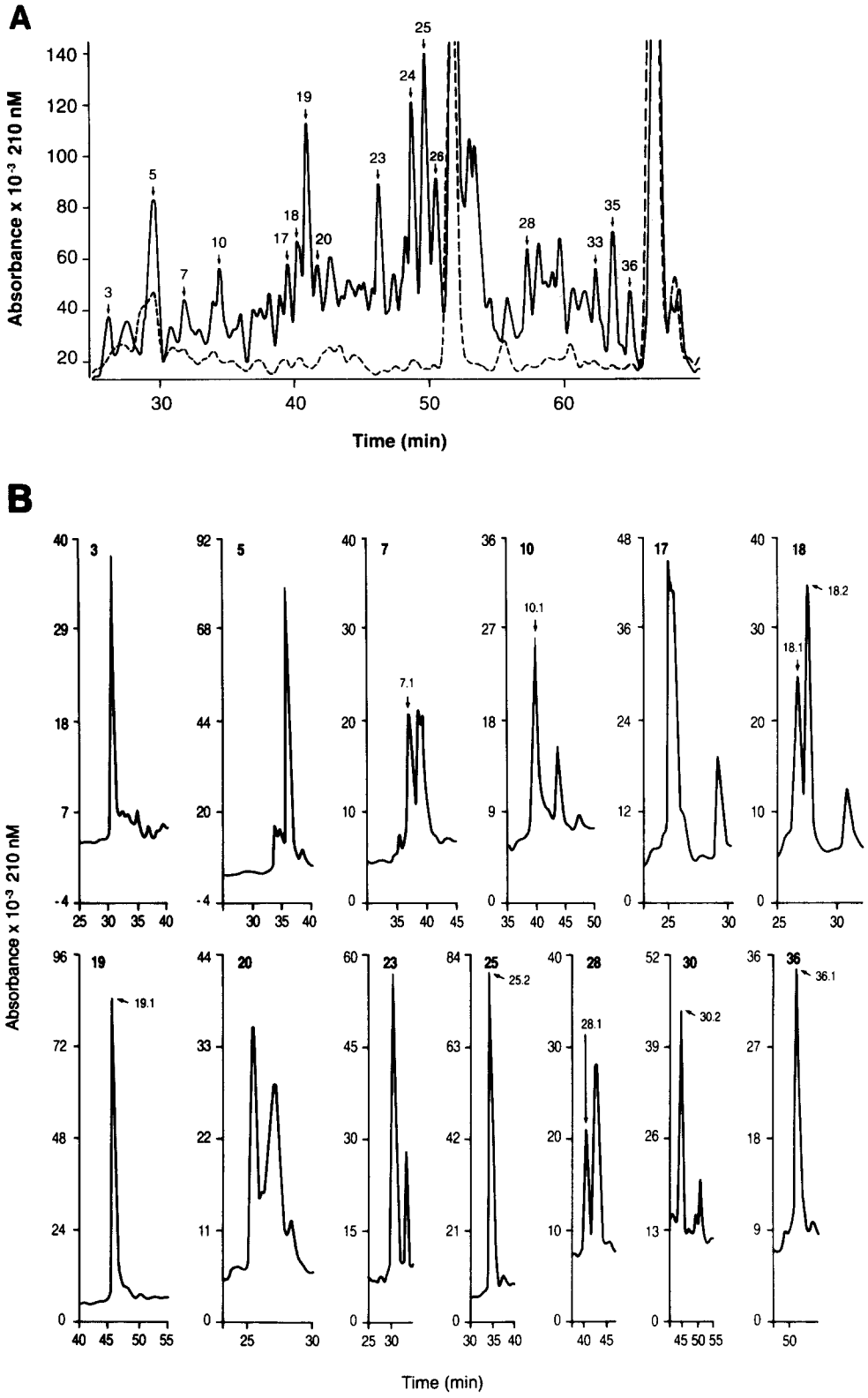

Figure 2. RP-HPLC chromatograms of peptides isolated from SQ7^b. (*A*) Separation of 90% of the total peptide fraction (Centricon 3 filtrate) isolated from affinity-purified SQ7^b molecules by acid denaturation followed by RP-HPLC (*solid line*). In the same experiment, peptides eluted from material purified from B22-249 column was also chromatographed (*broken line*). Chromatography was performed as described in Materials and Methods. (*B*) Chromatographs of repurification of individual peaks indicated by numbers in *A* by RP-HPLC under conditions different (see Materials and Methods) from the above. Peaks 24, 26, and 30 did not yield a detectable peak on repurification (data not shown). Note that the seemingly single peaks in the first chromatogram separate into two to three peaks. Peaks identified by numbers were subjected to automated amino acid sequencing by Edman method.

eluted peptide fraction) from the heavy and light chains by ultrafiltration.

To determine whether Qa-2 molecules bind peptides and, if they do, to determine the binding motif, 5% of the total acid-eluted peptide pool obtained from ~3.0 nmol of Qa-2 molecules (before HPLC fractionation) was subjected to 15 cycles of microsequence analysis. No significant phenylthiohydantoin (PTH) amino acid signals were detectable beyond

the ninth sequence cycle. Thus, the great majority of the peptides associated with SQ7^b are nonameric (Table 1). Most strikingly, these peptides have a single residue, histidine, at the seventh sequencing cycle (P7, position of the residue in the peptide) and a hydrophobic residue at the ninth cycle (Table 1). Further, only two to three different amino acid residues were detected at P1, P2, and P3, therefore, these positions also appear somewhat more conserved. In contrast, P4, P5, P6, and P8 are degenerate, i.e., a number of different structurally unrelated amino acids are represented at these positions (Table 1). These data suggest that the SQ7^b molecules bind peptides and have a specific binding motif.

Peptides Constitutively Bound to SQa-2 Are Derived from Intracellular Proteins. Sequencing of the pool of peptides bound to MHC molecules provides only limited information, such as the binding motif and the overall heterogeneity of the ligands. To gain further insight into the physico-chemical properties of the peptides bound to SQ7^b, such as the number of different peptides bound and sources of these peptides, we have purified peptides (~2.7 nmol of the acid-eluted fraction) eluted from SQ7^b molecules by RP-HPLC and sequenced a number of individual peaks. Approximately 50 distinct peaks were resolved by this method (Fig. 2 *A*, *solid line*) and 36 peaks were collected by hand. Due to the high background in the first separation and to ensure purity of each peak, 16 of the 36 peaks (Fig. 2 *A*, *arrows*) were further separated individually by RP-HPLC. In several instances, this procedure resulted in further separation of the single peaks into two or three peaks (Fig. 2 *B*).

A total of 18 individual sequences were derived from automated Edman sequencing of peaks indicated in Fig. 3 *B* because several of the peaks contained more than one peptide (Table 2). A comparison of the individual peptide sequences with the pool peptide data (Table 1) shows general agreement between them (Table 3). Almost all the individually sequenced peptides are nonameric (with the exception of peptide 10.1b [Table 2], which is a decamer). Further, the peptides predominantly use histidine at P7 (12/18 peptides) and a hydrophobic residue at the COOH terminus, either at P9 or at P10 (predominantly leucine; Table 2). The remaining positions, including the somewhat conserved P2 and P3, are occupied most frequently by the same residues that were detected in the pool peptide sequence data (Table 2). In addition, several amino acids that were not observed in the pool

Table 2. *Amino Acid Sequence of Peptides Constitutively Bound to SQ7^b*

Peptide peak ID	Position of residue 1	2	3	4	5	6	7	8	9	10	Protein source*
7.1 a	K	Q	N	P	I	A	H	Q	L	–	Unknown
b	A	L	A	E	L	P	H	E	I	–	Unknown
10.1 a	K	Q	N	P	T	V	H	H	L	–	Unknown
b	A	G	L	L	G‡	M	R	S	G	L§	Unknown
18.1 a	K	L	I	K	V	Y	H	S	L	–	Unknown
b	V	Q	N	X‖	T	M	H	P	I	–	Unknown
18.2	R	S	N	G	Q	V	H	M	L	–	Unknown
19.1 a	K	L	T	G	I	K	H	E	L	–	**Cofilin**
b	A	M	L	A	T	Y	H	K	L	–	Unknown
c	G	Q	L	X	V	X	H	K	L	–	Unknown
25.2 a	I	L	M	E	H	I	H	K	L	–	**Ribosomal L19**
b	A	Q	N	P	V	L	Y	Q	I	–	Unknown
c	S	M	I	K	T	V	R	E	M	–	Unknown
28.1	K	L	L	P	V	E	H	N	L	–	Unknown
30.2 a	K	Q	L	I	V	T	Y	H	L	–	Unknown
b	S	V	L	D	D	L	A	L	L	–	Unknown
c	Y	E	S	P	L	M	X	F	L	–	Unknown
36.1	D	L	L	G	T	L	H	N	L	–	Unknown

* Protein source was searched in the Swiss Protein Data Bank (SwissProt release 24,0; 1/93) using the Wisconsin Genetics Computer Group (GCG) sequence analysis package (27).

‡ Underline indicates probable presence of an alternate residue.

§ A 10th PTH amino acid was detected for that peptide; leucine could be the COOH-terminal residue in 10.1b peptide since P9 is represented by glycine for that particular peptide.

‖ X indicates that a PTH-amino acid could not be ascertained at that position.

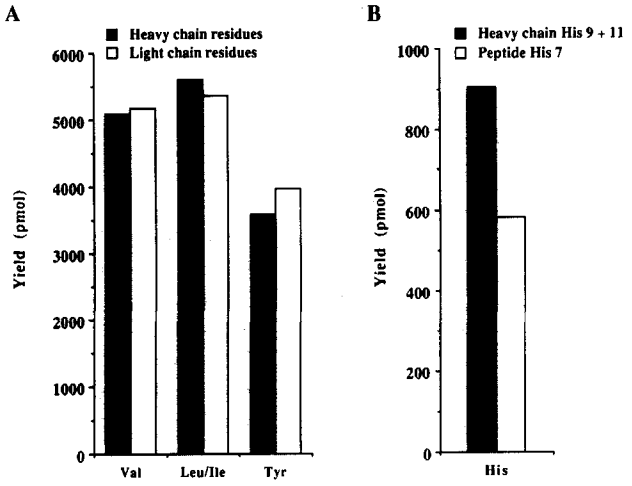

Figure 3. SQa-2 molecules exist as heterotrimers of heavy and light chains and peptides in approximately one stoichiometry. (*A*) Combined yields of amino acids in the types I and II SQ7^b heavy chains (*filled bars*) and β_2-m (*empty bars*) showing equimolar amounts of heavy and light chains. Val12, Leu5, and Tyr7 are residues of the type I heavy chain; Val14, Leu7, and Tyr9 are residues of the type II heavy chain; Val9, Ile7, and Tyr10 are residues of β_2-m used for this calculation. None of these residues are present in the peptide at the positions shown. (*B*) Combined yields (*filled bars*) of histidines at cycle 9 from the type I heavy chain and cycle 11 from the type II heavy chain compared with the histidine yield at cycle 7 from the peptide (*empty bars*).

peptide sequences were detected in the individual peptide sequences (Table 3). Most notable of these are arginine, tyrosine, and alanine at P7, and methionine and glycine at P9 (Tables 2 and 3). The higher degree of degeneracy observed at all positions in the individual peptide sequence data may be due to the differences in the sensitivity between the two approaches.

A homology search against a protein data bank revealed 2 of the 18 peptides, peptides 19.1a and 25.2a, had a 100% match with abundantly expressed murine intracellular proteins, cofilin (an actin-binding protein; 29) and L19 (a 60S ribosomal protein; 30) (Table 2). Both peptides represented major peaks in the RP-HPLC profile (Fig. 2 *A*) and were recovered with high yield (19.1a, 49.2 pmol; 25.2a, 110.7 pmol). None of the peptides have homology to serum proteins. This suggests that the SQa-2-associated peptides are derived from intracellular proteins of the host cell.

Native SQ7^b Is a Unimolar Complex of Heavy and Light Chains and Peptide. To approach the question of the stoichiometry of the Qa-2 heavy chain, light chain and peptides in the complex, we quantitated the affinity-purified SQ7^b molecules by sequence analysis. Amino acid sequencing of the first 20 residues of the complex revealed that two types of heavy chains are present in almost equimolar amounts in the complex (not shown). One type (I) begins with a glycine residue at the NH_2 terminus like almost all class I heavy chains (31), while the other type (II) has two additional amino acid residues (Arg1 and Ala2) at the NH_2 terminus, which is consistent with the previously reported data (32).

We have used the yields of residues that occur in the two heavy chains of SQ7^b and β_2-m to estimate the stoichiometry of the heavy and light chains in the class I complex. The choices of the residues used as markers for this analysis were based on the following criteria. (*a*) When identical residues were compared, they occurred in the heavy chain and β_2-m at different sequencing cycles; however, to have comparable yields, the residues had to occur at close posi-

Table 3. *Compilation of Different Amino Acid Residues that Occur at a Given Position in Individual Peptides and Comparison to Pool Peptide Sequences*

Position in the peptide									
P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
K*(6)‡	**Q** (6)	**L** (7)	**P** (5)	**I** (2)	**K** (1)	**H** (12)	**E** (3)	**L** (13)	L (1)
A (4)	**L** (6)	**N** (5)	**E** (2)	**V** (5)	**L** (3)	R (2)	**Q** (2)	**I** (3)	
G (1)	M (2)	I (2)	**G** (3)	**Q** (1)	**F** (0)	Y (2)	**N** (2)	**F** (0)	
S (2)	G (1)	A (1)	**K** (2)	**E** (0)	**M** (3)	A (1)	**K** (3)	G (1)	
I (1)	S (1)	T (1)	L (1)	**L** (2)	**Y** (2)		**S** (2)	M (1)	
D (1)	V (1)	M (1)	A (1)	**T** (5)	V (3)		**T** (0)		
R (1)	E (1)	S (1)	I (1)	**Y** (0)	A (1)		**V** (0)		
V (1)			D (1)	G (1)	P (1)		H (2)		
Y (1)				H (1)	I (1)		L (1)		
				D (1)	E (1)		F (1)		
					T (1)		M (1)		
							P (1)		

* Bold letters indicate amino acid assignments from pool peptide sequences (see Table 1).
‡ Number of times a residue appears at a given position on compiling the individual peptide sequences (see Table 2) is given in parenthesis.

tions in the heavy and the light chains. (*b*) When two different residues were compared, they occurred in the same sequencing cycle, and their relative yields, in general, are known to be very similar, e.g., leucine and isoleucine. Thus, the combined yields of the two heavy chains using residues Val12, Leu5, and Tyr7 for type I and Val14, Lev7, and Tyr9 for type II heavy chains were compared to the yields of residues Val9, Ile7, and Tyr10 of β_2-m (Fig. 3 *A*). The data demonstrated that the heavy chains are associated with the light chain (β_2-m) in unimolar ratio (Fig. 3 *A*). Further, both the mouse and bovine β_2-m are present in the complex in similar amounts as judged by the yields of their specific residues (data not shown).

Amino acid sequence analysis was also used to quantitate the peptides bound to the $SQ7^b$ molecules. For this quantitation, histidine was chosen as the marker residue because of the fact that it is the most common amino acid in the peptide at P7 (Table 2) and is not represented in the heavy or light chains at cycle 7. Histidine is present at cycle 9 in type I and at cycle 11 in type II $SQ7^b$ heavy chains (32) but not in the peptide at these cycles (this report). This comparison revealed that P7His represented ~65% of the yield of histidines at cycles 9 and 11 of the two heavy chain sequences (Fig. 3 *B*). Since histidine at P7 occurred in 12 of 18 peptides sequenced, and if this distribution reflects the average occurrence of this residue in the peptide, one would estimate peptide occupancy of ~97%.

In another experiment designed to address this issue, we have directly determined the molar amounts of the purified $SQ7^b$ molecules and the acid-eluted peptide fraction. To make this calculation, 2.5% of the starting amount of the affinity-purified heterotrimeric complex and 10% of the acid-eluted peptide fraction were quantitated by amino acid analysis. Using the yields of alanine, leucine, and arginine, it was determined that the affinity-purified starting material contained 5.75 nmol SQa-2 protein (see Materials and Methods). Similarly, using the yields of the conserved histidine and leucine residues in the peptides, it was determined that the acid-eluted ultrafiltrate contained ~53% molar yield of peptides expected from 5.75 nmol of $SQ7^b$ molecules (see Materials and Methods) if there were 1:1 peptide/MHC stoichiometry. This is in good agreement with the value of ~65% obtained from the sequence analysis of the trimolecular complex. Both these values represent lower estimates because histidine is not the only residue present at P7 in all the peptides. In fact, as mentioned above, since histidine is represented only in ~65% of the individually sequenced peptides (Table 2), when corrected the actual value would be ~95%. Together, the data are consistent with the conclusion that the majority of the SQa-2 molecules are occupied by peptides at least to the extent seen with yields for HLA-B27-associated peptides (33). Thus, native $SQ7^b$ molecules appear to be predominantly heterotrimers consisting of noncovalently associated heavy and light chains complexed with peptides in close to 1:1:1 stoichiometry.

Several Hundred Different Peptides Are Constitutively Bound to SQa-2 Molecules. The properties of the peptides associated with $SQ7^b$ suggest that they are very similar to those bound to class Ia molecules. Since individual class Ia molecules can bind a large variety of peptides of different but related sequences, somewhere in the range of 200 to >1,000 (34), we addressed the question of the size of the peptide repertoire that could associate with $SQ7^b$ molecules. One estimate could be made by directly counting the number of peaks obtained from the two RP-HPLC chromatograms and the number of peptide sequences obtained per peak. The number of peptide peaks from the first RP-HPLC separation is ~50 (Fig. 2 *A*). Several of these peaks on rechromatography yielded an average of two peaks (Fig. 2 *B*), making the total number of resolvable peaks to be in the range of 100. Since, an average of two peptide sequences were obtained per peak (Table 2), it is reasonable to conclude that at least 200 diverse peptides can bind to $SQ7^b$ molecules. This would be a minimum estimate of the size of the peptide repertoire due to the sensitivity of the analysis, which does not detect minor species of peptides that may be associated with the class I molecules.

The above estimate is consistent with another calculation based on molar yields of the total peptides in the Centricon 3 ultrafiltrate eluted from $SQ7^b$ and the yields of individual peptides. If 3.03 nmol of the starting amount of peptides in the ultrafiltrate contains 100 different peptides, then a single peptide that occupies 1.0% of $SQ7^b$ should have a yield of 30.3 pmol. Since the majority (15/18) of the peptides sequenced were represented in amounts <30.3 pmol (range: 2.0–22.6 pmol, 0.07–0.7%; exceptions: peptide 19.1a, 1.6%; 25.2a, 3.6%; 25.2b, 1.4%), we conclude that $SQ7^b$ molecules can bind >100 different peptides.

Discussion

The availability of SQa-2 molecules in large amounts produced by genetically engineered mammalian cells allowed a thorough biochemical analysis of this molecule using sensitive techniques. Our data demonstrate that $SQ7^b$, an MHC class Ib molecule, constitutively binds several hundred different naturally processed peptides. In most respects, the properties of the bound peptides, such as the nonameric length, the presence of conserved residues at two positions, the intracellular source of the peptides, and the minimum estimated number of different peptides that can bind Qa-2 molecules, are similar to MHC class Ia binding peptides.

A more refined understanding of the properties of the $SQ7^b$ binding peptides was achieved by examining the individual peptide sequences than was possible from pool sequence data. The most striking similarity between the data obtained from the pool sequence and that of individual peptides is the restricted usage of histidine at P7 and the presence of a hydrophobic residue at the carboxy terminus (PC). Interestingly, but less frequently, arginine, tyrosine, and alanine were seen at P7. These residues were not represented in the pool sequences probably because the relative molar abundance (~20% when the yields of the three residues are combined) of these amino acids may have been below the threshold of detection under the conditions used for pool sequencing. In contrast to the conserved amino acid usage at P1, P2, and P3 deduced from the peptide pool sequencing, these posi-

tions show high sequence variability on analysis of individual sequences. Whereas P1 is genuinely structurally variable, P2 and P3 show a propensity for either a hydrophobic or a polar residue. Thus, P2 and P3 are probably moderately conserved positions. The remainder of the positions (P4, P5, P6, and P8) in the peptide are degenerate. Thus, the data for individual peptide sequences point out that the mechanism for peptide MHC interaction is complex and simply knowing the allele specific-motif is not sufficient to predict the precise rules for peptide binding.

Based on the structural motif of the $SQ7^b$ binding peptides, we suggest that the side chains of residues at P7 and PC (P9/P10) probably serve as specificity anchors for binding to $SQ7^b$ cleft. In addition to the dominant specificity anchors, side chains of residues at P2 and P3 may play the role of accessory or alternate anchors that may be required for the specificity and/or affinity for binding to $SQ7^b$. It should be noted that at least one of the conserved amino acid residues (P2, Gln or Leu; P3, Leu or Asn), or a structurally related residue, is used at P2 and/or P3 in the peptides that lack the dominant P7His anchor (Table 2). Thus, in such peptides, P2 and/or P3 may play an important role in anchoring the peptide. On average, individual peptides contain only three of the four proposed anchors at P2, P3, P7, and P9. This is similar to the number of anchors observed in HLA-A3 binding peptides (35). A unique feature of the $SQ7^b$ binding motif is the fact that, besides the COOH-terminal anchor (PC), the dominant anchor at P7 of the peptide is located distally towards the COOH-terminal end of the peptide. This is in striking contrast with all the known allele-specific dominant anchor positions of mouse (36–38) and human (33–35, 39, 40) class Ia binding peptides and may be the consequence of the physico-chemical architecture of the Qa-2 groove.

Rotzschke et al. (41) have recently reported the sequence of peptide pool eluted from GPI-linked splenic and thymic Qa-2 complexes and proposed that "only a few different peptides are presented by these molecules" based on the assumption that each peptide uses four to five anchors for binding Qa-2. Furthermore, they suggested that Qa-2 functions as a restriction element for peptides derived from gut-associated antigens. Whereas the sequence data reported for the peptide pool presented here and by Rotzschke et al. (41) are extremely similar, the two studies differ drastically in the qualitative and quantitative estimates of Qa-2 binding peptide repertoire. Using two approaches, we have determined empirically that a minimum of 200 different peptides can bind $SQ7^b$, a number similar to the minimum estimates of class Ia binding peptide repertoire. Several other observations also support our conclusion. First, our data from individual peptide sequences demonstrate that only a subset of Qa-2 binding anchors at P2, P3, P7, and P9 is present in individual peptides. This, we suggest, enables the binding of a large array of diverse peptides to Qa-2. Second, the function of the Qa-2 antigens in the immune response against gut-associated bacteria is inconsistent with the known tissue distribution of these molecules (expressed only in lymphoid derived cells; reference 3). Our computer searches of data banks using anchors at P7 and P9 together with auxillary anchors at either P2 or P3 did not reveal homology to any particular group of proteins, pathogen derived or otherwise (Stroynowski, I., and B. Jameson, unpublished results). Third, if the criteria used by Rotzschke et al. (41) to define Qa-2 binding motif in the peptides were indeed correct, then their computer searches should have found peptides derived from cofilin and L19 (reported here), both of which are mouse proteins (29, 30) found in SwissProt, GenBank, and EMBL sequence data banks. Finally, the ability to express SQa-2 and GPI-Qa-2 in high amounts in a variety of transfected cells, including myeloma (this report), hepatomas, and L cells (18), and CHO cells (Tabaczewski, P., and I. Stroynowski, unpublished results). also argues against the contention that Qa-2 binding peptides are rare or limiting. Thus, Qa-2 molecules are not "receptors of higher stringency than ordinary class I molecules (41)" but they bind a large array of diverse peptides derived from endogenously synthesized and naturally processed proteins in much the same manner as the ordinary class I molecules.

We gratefully acknowledge the technical help provided by Mr. Yuan Shi and Ms. Nguyet Le for the amino acid sequencing and analysis in a timely manner. We thank Dr. G. Kepecs for help with the HPLC, Dr. E. C. Goyarts for sharing sequence information of $Q7^b$ cDNA, Drs. G. Kepecs and N. Papadapaulos for critical reading of the manuscript, and Ms. R. Spata for the preparation of the manuscript.

This work was supported by a Cancer Research Institute postdoctoral fellowship (S. Joyce) and grants from the National Institutes of Health (S. G. Nathenson and I. Stoynowski).

Address correspondence to Stanley G. Nathenson, Department of Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461 and Iwona Stroynowski, Department of Microbiology and the Center for Diabetes Research, University of Texas Southwestern Medical Center at Dallas, 53523 Harry Hines Boulevard, Dallas, TX 75235-8854.

Received for publication 5 October 1993.

References

1. Townsend, A.R.M., and H. Bodmer. 1989. Antigen recognition by class I restricted T lymphocytes. *Annu. Rev. Immunol.* 7:601.
2. Zinkernagel, R.M., and P.C. Doherty. 1979. MHC-restricted cytotoxic T cells: Studies on the biological role of polymorphic major transplantation antigens determining T-cell restriction-specificity, function, and responsiveness. *Adv. Immunol.* 27:51.
3. Stroynowski, I. 1990. Molecules related to the class I major histocompatibility complex antigens. *Annu. Rev. Immunol.* 8: 501.
4. Flaherty, L., E. Elliott, J.J. Tine, A. Walsh, and J. Waters. 1990. Immunogenetics of the Q and *TL* regions of the mouse. CRC *Crit. Rev. Immunol.* 10:131.
5. Fischer Lindahl, K., E. Hermel, B.E. Loveland, and C.R. Wang. 1991. Maternally transmitted antigen of mice: a model transplantation antigen. *Annu. Rev. Immunol.* 9:351.
6. Van Kaer, L., M. Wu, Y. Ichikawa, K. Ito, M. Bonneville, S. Ostrand-Rosenberg, D.B. Murphy, and S. Tonegawa. 1991. Recognition of MHC *TL* gene products by $\gamma\delta$ T cells. *Immunol. Rev.* 120:89.
7. Kurlander, R.G., S.M. Shawar, M.L. Brown, and R.R. Rich. 1992. Specialized role for a murine class Ib MHC molecule in prokaryotic host defenses. *Science (Wash. DC).* 257:678.
8. Palmer, E.G., C.-R. Wang, L. Flaherty, K. Fischer Lindahl, and M.J. Bevan. 1992. H-2M3 presents Listeria monocytogenes peptide to cytotoxic T lymphocytes. *Cell.* 70:215.
9. Milligan, G.N., L. Flahergy, V.L. Braciale, and T.J. Braciale. 1991. Nonconventional (*TL* encoded) major histocompatibility complex molecules present processed viral antigen to cytotoxic T lymphocytes. *J. Exp. Med.* 174:133.
10. Cheseboro, B., M. Miyazawa, and W.J. Britt. 1990. Host genetic control of spontaneous and induced immunity to Friend murine retrovirus infection. *Annu. Rev. Immunol.* 8:477.
11. Miyazawa, M., J. Nishio, K. Wehrly, C.S. David, and B. Cheseboro. 1992. Spontaneous recovery from Friend retrovirus induced leukemia. Mapping of the *Rfv-2* gene in the *Q/TL* region of the mouse MHC. *J. Immunol.* 148:1964.
12. Loveland, B., C.-R. Wang, H. Yonekawa, E. Hermel, and K. Fischer-Lindahl. 1990. Maternally transmitted histocompatibility antigen of mice: a hydrophobic peptide of a mitochondrially encoded protein. *Cell.* 60:971.
13. Shawar, S.M., R.G. Cook, J.R. Rodgers, and R.R. Rich. 1990. Specialized functions of MHC class I molecules: I. an *N*-formyl peptide receptor is required for the construction of class I antigen Mta. *J. Exp. Med.* 171:897.
14. Shawar, S.M., J.M. Vyas, J.R. Rodgers, R.G. Cook, and R.R. Rich. 1991. Specialized functions of MHC class I molecules. II. Hmt binds *N*-formylated peptides of mitochondrial and prokaryotic origin. *J. Exp. Med.* 174:941.
15. Wang, C.-R., B.E. Loveland, and K. Fischer Lindahl. 1991. *H-2M3* encodes the MHC class I molecule presenting the maternally transmitted antigen of the mouse. *Cell.* 66:335.
16. Vyas, J.M., S.M. Shawar, J.R. Rodgers, R.G. Cook, and R.R. Rich. 1992. Biochemical specificity of H-2M3^s: sterospecificity and space-filling requirements at position 1 maintain N-formyl peptide binding. *J. Immunol.* 149:3605.
17. Soloski, M.J., L. Hood, and I. Stroynowski. 1988. *Qa* region class I gene expression: identification of a second class I gene, *Q9*, encoding a Qa-2 polypeptide. *Proc. Natl. Acad. Sci. USA.* 85:3100.
18. Stroynowski, I., M. Soloski, M. Low, and L. Hood. 1987. A single gene encodes soluble and membrane-bound forms of the major histocompatibility Qa-2 antigen: anchoring of the product by a phospholipid tail. *Cell.* 50:759.
19. Waneck, G.L., D.H. Sherman, P.W. Kincade, M.G. Low, and R.A. Flavell. 1988. Molecular mapping of signals in the Qa-2 antigen required for attachment of the phosphotidylinositol membrane anchor. *Proc. Natl. Acad. Sci. USA.* 85:577.
20. Ulker, N., K.D. Lewis, L. Hood, and I. Stroynowski. 1990. Activated T cells transcribe an alternatively spliced mRNA encoding a soluble form of Qa-2 antigen. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:3839.

20a. Tabaczewski, P., H. Shirwan, K. Lewis, and I. Stroynowski. 1994. Alternative splicing of class Ib MHC transcripts *in vivo* leads to the expression of soluble Qa-2 molecules in murine blood. *Proc. Natl. Acad. Sci. USA.* 91. In press.

21. Hendricks, M.B., M.J. Banker, and M. McLaughlin. 1988. A high efficiency vector for expression of foreign genes in myeloma cells. *Gene (Amst.).* 64:43.
22. Devlin, J.J., E.H. Weiss, M. Paulson, and R.A. Flavell. 1985. Duplicated gene pairs and alleles of class I genes in the Qa-2 region of murine major histocompatibility complex: a comparison. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:3203.
23. Harlow, E., and D. Lane. 1988. Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
24. Hasenkrug, K.J., J.M. Cory, and J.H. Stimpfling. 1987. Monoclonal antibodies defining mouse tissue antigens encoded by the *H-2* region. *Immunogenetics.* 25:136.
25. Sharrow, S.O., J.S. Arn, I. Stroynowsi, L. Hood, and D.H. Sachs. 1989. Epitope clusters of Qa-2 antigens defined by a panel of new monoclonal antibodies. *J. Immunol.* 142:3495.
26. Allen, H., D. Wraith, D. Pala, B. Askonas, and R.A. Flavell. 1984. Domain interactions of H-2 class I antigens alter cytotoxic T cell recognition sites. *Nature (Lond.).* 309:279.
27. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387.
28. Tian, H., F. Imani, and M.J. Soloski. 1991. Physical and molecular genetic analysis of Qa-2 antigen expression: Multiple factors controlling cell surface levels. *Mol. Immunol.* 28:845.
29. Moriyama, K., S. Matsumoto, E. Nishida, H. Sakai, and I. Yahara. 1990. Nucleotide sequence of the mouse cofilin cDNA. *Nucleic Acids. Res.* 18:3053.
30. Nakamura, T., M. Onno, R. Mariage-Samson, J. Hillova, and M. Hill. 1990. Nucleotide sequence of mouse L19 ribosomal protein cDNA isolated in screening with *tre* oncogene probes. *DNA Cell Biol.* 9:697.
31. Kabat, E.A., T.T. Wu, H.M. Perry, K.S. Gottesman, and C. Foeller. 1991. Sequences of Proteins of Immunological Interest. National Institutes of Health Publication No. 91-3242. U.S. Department of Health and Human Services, Bethesda, MD.
32. Widacki, S.M., and R.G. Cook. 1989. NH2-terminal amino acid sequence analysis of Qa-2 alloantigens. *Immunogenetics.* 29:206.
33. Jardetzky, T.S., W.S. Lane, R.A. Robinson, D.R. Madden, and D.C. Wiley. 1991. Identification of self peptides bound to purified HLA-B27. *Nature (Lond.).* 353:326.
34. Hunt, D.F., R.A. Henderson, J. Shabanowitz, K. Sakaguchi, H. Michel, N. Seviler, A.L. Cox, E. Appella, and V.H. Engelhard. 1992. Characterization of peptides bound to the class I MHC molecules HLA-A2.1 by mass spectrometry. *Science*

(Wash. DC). 255:1261.

35. DiBrino, M., K.C. Parker, J. Shiloach, M. Knierman, J. Lukszo, R.V. Turner, W.E. Biddison, and J.E. Colligan. 1993. Endogenous peptides bound to HLA-A3 possess a specific combination of anchor residues that permit identification of potential antigenic peptides. *Proc. Natl. Acad. Sci. USA.* 90:1508.
36. Van Bleek, G.M., and S.G. Nathenson. 1991. The structure of the antigen binding groove of major histocompatibility complex class I molecules determines specific selection of peptides. *Proc. Natl. Acad. Sci. USA.* 88:11032.
37. Falk, K., O. Rotzschke, S. Stevanovic, G. Jung, and H.-G. Rammensee. 1991. Allele specific motifs revealed by sequencing of self peptides eluted from MHC molecules. *Nature (Lond.).* 351:290.
38. Corr, M., L.F. Boyd, S.R. Frankel, S. Kozlowski, E.A. Padlan, and D.H. Margulies. 1992. Endogenous peptides of a soluble major histocompatibility complex class I molecule, H-2L^d_s: sequence motifs, quantitative binding, and molecular modeling of the complex. *J. Exp. Med.* 176:1681.
39. Guo, H.-C., T.S. Jardetzky, T.P.J. Garrett, W.S. Lane, J.L. Strominger, and D.C. Wiley. 1992. Different length peptides bind to HLA-Aw68 similarly at the ends but bulge out in the middle. *Nature (Lond.).* 360:364.
40. Sutton, J., S. Rowland-Jones, W. Rosenberg, D. Nixon, F. Gotch, X.-M. Gao, N. Murray, A. Spoonas, P. Driscoll, M. Smith, A. Willis, and A. McMichael. 1993. A sequence pattern for peptides presented to cytotoxic T lymphocytes by HLA-B8 revealed by analysis of epitopes and eluted peptides. *Eur. J. Immunol.* 23:447.
41. Rotzschke, O., K. Falk, S. Stevanovic, B. Grahovac, M.J. Soloski, G. Jung, and H.-G. Rammensee. 1993. Qa-2 molecules are peptide receptors of higher stringency than ordinary class I molecules. *Nature (Lond.).* 361:642.