Endogenous Peptides of a Soluble Major Histocompatibility Complex Class I Molecule, H-2L^d_s: Sequence Motif, Quantitative Binding, and Molecular Modeling of the Complex

By Maripat Corr, Lisa F. Boyd, Shlomit R. Frankel, Steven Kozlowski, Eduardo A. Padlan,* and David H. Margulies

From the Molecular Biology Section, Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, and the *Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Summary

To gain insight into the rules that govern the binding of endogenous and viral peptides to a given major histocompatibility complex (MHC) class I molecule, we characterized the amino acid sequences of a set of self peptides bound by a soluble analogue of murine H-2L^d, H-2L^d_s. We tested corresponding synthetic peptides quantitatively for binding in several different assays, and built three-dimensional computer models of eight peptide/H-2L^d_s complexes, based on the crystallographic structure of the human HLA-B27/peptide complex. Comparison of primary and tertiary structures of bound self and antigenic peptides revealed that residues 2 and 9 were not only restricted in sequence and tolerant of conservative substitutions, but were spatially constrained in the three-dimensional models. The degree of sequence variability of specific residues in MHC-restricted peptides reflected the lack of structural constraint on those amino acids. Thus, amino acid residues that define a peptide motif represent side chains required or preferred for a close fit with the MHC class I heavy chain.

The cell surface antigens encoded by the MHC class I genes consist of a polymorphic H chain of 45 kD bound to a peptide of 8–10 residues and a noncovalently bound 12-kD L chain, β_2 microglobulin (β_2-m)[1] (1, 2). The bound peptides usually derive from the MHC-expressing cell and represent proteolytic fragments of cytosolic, plasma membrane-associated, or secreted proteins (3, 4). Upon infection with cellular parasites such as viruses or bacteria, specific peptide fragments can be bound by the class I molecule and displayed at the cell surface (5–7). There, the MHC class I/peptide/β_2-m complex serves as a major element in the recognition of such infected cells by clonotypic TCRs (1, 8, 9). Detailed x-ray structures of the MHC class I molecules HLA-A2 (10–12), HLA-A68 (13), HLA-B27 (14), and H-2K^b (15, 16) have been reported which reveal the intimate association of self peptide with a groove supported by antiparallel β-sheet and bounded by the helices of the α_1 and α_2 domains.

Several approaches have been employed to evaluate the nature of self and antigenic peptides bound to MHC proteins.

First, the elution of peptides with acid, the isolation of a single pool of peptides by reverse phase HPLC, and the determination of the preferred amino acids at particular positions by automated Edman degradation, have led to the definition of particular motifs for peptides bound to certain MHC molecules (2, 6, 17). Second, individual peptides have been identified by the sequencing of single HPLC fractions, using either classical techniques or mass spectrometry (3–5). In one case, the identification of the peptide sequence motifs was the basis of the molecular interpretation of electron density maps of the HLA-B27 molecule (14). Detailed evaluation of the ability of self peptides to bind the MHC H chain/β_2-m complex from which they were eluted and the correlation of this binding with structural models has been limited (4, 18). We report here the identification of self peptides eluted from a soluble analogue of H-2L^d, the evaluation of their binding to both the soluble H-2L^d as well as cell surface H-2L^d, and molecular modeling of the similarities and differences of eight H-2L^d/peptide/β_2-m complexes.

Materials and Methods

Protein and Synthetic Peptides. The soluble analogue of H-2L^d,

---

[1] *Abbreviations used in this paper:* β_2-m, β_2 microglobulin; LCMV, lymphocytic choriomeningitis virus; MCMV, murine cytomegalovirus; NHS, N-hydroxysuccinimide; RU, resonance unit.

H-2L^d,, is a chimeric molecule expressed as the α_1 and α_2 domains of H-2L^d with the α_3 domain and COOH terminus of Q10^b (19). Protein was immunoaffinity purified from transfectant L cell culture supernatants on a 30-5-7S (20)-coupled Sepharose 4B column and concentrated by centrifugation (Centricon 10; Amicon Corp., Beverly, MA) (21). The mAb 30-5-7S is specific for the α_2 domain of maturely folded H-2L^d molecules (22-24). Synthetic peptide corresponding to the H-2L^d restricted murine cytomegalovirus (MCMV) pp89 early regulatory protein residues 168-176 (YPHFMPTNL), pMCMV, was purchased from Synthecell Corp. (Rockville, MD) or provided by John Coligan of the Biological Resources Branch (BRB), National Institute of Allergy and Infectious Diseases (NIAID). All other peptides (provided by J. Coligan) were synthesized using standard Fmoc chemistry on peptide synthesizers, repurified by reverse phase HPLC, and sequence confirmed on an automated sequenator. A panel of synthetic peptides was patterned after the sequences identified for the acid-eluted H-2L^d, self peptides with alanine residues inserted at positions where the sequence was not clearly defined. Each synthetic peptide was designated by the HPLC fraction from which the sequence was originally identified. A peptide, pAPAL, containing the H-2L^d core motif, APA$_6$L, was also synthesized.

Peptide Elution and Sequence Analysis. 435 μg of purified H-2L^d, were brought to 0.1% TFA at ambient temperature and separated by reverse phase HPLC using a C18 column (Vydac, Hesperia, CA [4.6 × 250 mm, 5-μM particle size]). The sample was loaded in 0.01% TFA in dH$_2$O for 5 min, and then eluted with a 0-60% linear acetonitrile gradient over 35 min as described (2). After the gradient, the remaining material was eluted with 10 min of 80% acetonitrile. The flow rate was 1 ml/min and the absorbance was detected at 220 nm. 1-ml fractions were collected, dried by vacuum centrifugation, and recovered fractions 21-30 were sequenced by automated Edman degradation on a sequenator (model 470A; Applied Biosystems, Inc., Foster City, CA), standardized to a repetitive yield of 94%. Individual fractions were sequenced so as not to overlook peptides that occurred at low abundance. Residues were identified by amino acid analysis on a model 120A PTH analyzer (Applied Biosystems, Inc., Foster City, CA). All sequencing and amino acid analysis was performed by the BRB, NIAID (Frederick, MD). Homologous sequences were identified by searching the National Biomedical Research Foundation (NBRF) data bank using MacVector v.3.5 (International Biotechnologies Inc., New Haven, CT).

Competitive Binding Analysis. Synthetic pMCMV was radiolabeled at its NH$_2$-terminal tyrosine residue with ^{125}I and purified by BioGel (Bio-Rad Laboratories, Richmond, CA) P-2 gel filtration and reverse phase HPLC as described (21). Specific activity of the labeled peptide was 8.9 × 10^{15} dpm/mole for the experiments shown in Fig. 2, *A* and *B*, and 9.6 × 10^{15} dpm/mole in Fig. 2 *C*. Affinity-purified H-2L^d, (800 nM) was added to a mixture of 2 μM radiolabeled peptide and varying amounts of cold competing peptides in 0.02% BSA in PBS. Samples were placed in siliconized, blocked (with 2% BSA [Grade V; Sigma Chemical Co., St. Louis, MO] in 0.02% azide) and rinsed microfuge tubes, overlayed with light mineral oil, and incubated at 37°C for 45-47 h. Analysis of equilibrium binding was done by spin-column chromatography as described (21). For competition binding analysis, bound and free counts for experimental and nonspecific control samples were analyzed with LIGAND (Mac version 4.1) (25) and ALLFIT (Mac version 1.0) (26).

Serological Detection Using Surface Plasmon Resonance. The binding of mAb 30-5-7S to various H-2L^d,/peptide complexes was assessed by surface plasmon resonance using BIAcore™ (Pharmacia

Biosensor, Piscataway, NJ). Ester groups of the carboxymethylated dextran matrix of the gold surface sensor chip CM5™ (Pharmacia Biosensor AB, Uppsala, Sweden) were chemically activated during a 6-min continuous flow (5 μl/min) exposure to 50 mM N-hydroxysuccinimide (NHS) and 200 mM N-ethyl-N'-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC), and then covalently linked to the free amines of 30-5-7S (0.02 mg/ml in 10 mM Na acetate pH 5.0) for 6 min. The remaining reactive groups were then blocked with 1 M ethanolamine hydrochloride, pH 8.5. The integrated microfluidic cartridge (IFC) was thermally regulated at a constant temperature of 25°C, and the photodiode array detection signals were normalized with 40% wt/wt glycerol. All solutions were filtered and degassed. 50 μl of 10 μg/ml immunoaffinity purified H-2L^d, (in 10 mM Hepes, pH 7.4, 150 mM NaCl, 3.4 mM EDTA, and 0.05% BIAcore Surfactant P20 (Pharmacia Biosensor) (Hepes buffered saline [HBS]) were added to 50 μl of varying amounts of peptide in HBS or HBS alone at room temperature to achieve equilibrium (at least 15 h). 30-μl aliquots were analyzed for direct binding to the immobilized 30-5-7S on the matrix surface by surface plasmon resonance (27-31). The flow rate was constant at 5 μl/min and the antibody surface was regenerated with a 1-min wash of 5 mM phosphoric acid. 1,000 resonance units (RU) correspond to a surface concentration of $\sim$1 ng/mm^2. Normalized RUs provided a quantitative indication of the mass of H-2L^d, associated with the antibody at a given time.

Analysis of Sequence Variability. Peptide sequences were analyzed for sequence variability according to the algorithm of Wu and Kabat (32), and for structural variability according to Padlan (33) using the amino acid dissimilarity values of Grantham (34).

Molecular Modeling. Three-dimensional coordinates for the HLA-B27/peptide complex were kindly provided by D. Madden and D. Wiley (Harvard University, Cambridge, MA) (14). The amino acid structures corresponding to the H-2L^d sequence for residues 1-182 (35-38), the Q10^b sequence for residues 183-275 (39), and for murine β_2-m (40, 41) were substituted at the positions of the corresponding residues in the HLA-B27 structure using the computer graphics program FRODO (42). Alpha carbons, C$_\alpha$, of the substituted amino acids were superimposed on the position of the C$_\alpha$ of the corresponding residue of HLA-B27. There was no need to introduce gaps in either sequence to maximize the homology alignment. Similarly, antigenic peptide residues were substituted onto the corresponding positions of the HLA-B27 bound peptide. Prolines were substituted without difficulty in the *trans* configuration in all cases. After the substitution, the amino acid side chains were regularized to eliminate geometric incompatibilities, and the resulting coordinates were subjected to 40 steps of minimization using the Powell algorithm (43) in the program X-PLOR (44). These structures were subjected to simulated annealing (45, 46) using X-PLOR by heating to 1,000° K and slow cooling to 300° K. The only constraints in the simulated annealing were the disulfide bonds joining residues 101-164 and 203-259 of the H chain and residues 25-80 of the β_2-m L chain. The structure was again energy minimized through 120 steps of Powell minimization, and then subjected to minimization by the adopted basis Newton-Raphson technique (47) through an additional 100 cycles. Final coordinates for each structure were compared for root mean square (rms) deviations and molecular surface exposure. Connolly (48) surfaces were calculated for a water probe of 1.7 Å . Calculations were performed on VAX 8530 (Powell minimization and simulated annealing) and IRIS 4D/70G (Silicon Graphics Inc., Mountain View, CA) (ABNR minimization and graphics) (using QUANTA 3.2.3; Polygen Corp., Waltham, MA). The validity of the simulated annealing procedure in building these models is emphasized by the

demonstration that a similar model of the HLA-A2 molecule compared favorably with the crystal structure of HLA-A2 with rms deviation of all 275 C_α atoms of < 1.5 Å. Atomic coordinates for the models described here will be made available on request.

Results

Our laboratory has demonstrated the biological function and binding of a viral antigenic peptide to a soluble analogue of the murine class I molecule H-2L^d, H-2L^d_s (21), and we were interested in determining the sequence of self peptides that copurify with this molecule. H-2L^d_s was purified from transfected L cell culture supernatants by immunoadsorbent chromatography using a mAb (30-5-7S) that recognizes a conformationally sensitive epitope found on mature, peptide-complexed H-2L^d molecules (49, 50) (Corr, M. and D. Margulies, unpublished results). Over 400 μg of purified H-2L^d_s were dissociated by treatment with acid, and the components of the class I ternary complex were separated by reverse phase HPLC under conditions such that most peptides of 8–11 residues in length would be eluted with retention times between 20 and 29 min (2) (Fig. 1). Individual fractions (21–30), containing small pools of peptides, were subjected to sequence analysis using automated Edman degradation (Table 1). Frequently, the amino acid analysis of the first Edman step was heterogeneous, so assignments at the first position were difficult. Fractions 21 and 22 failed to yield reliable sequence information. Also, the PTH amino acid yields for fractions
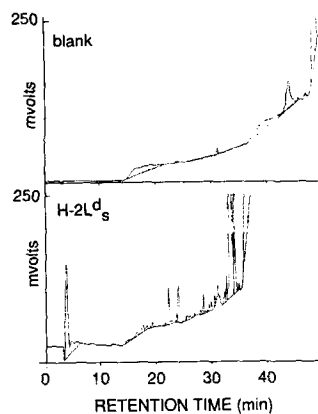

Figure 1. Reverse phase HPLC profile of H-2L^d_s molecules. H-2L^d_s molecules were immuno-affinity purified, denatured with 0.1% TFA, and loaded onto a C18 reverse phase HPLC column as described in Materials and Methods. Peptides were then eluted with a 0–60% linear acetonitrile gradient between 5 and 40 min. (Top) mock injection; (bottom) H-2L^d_s/peptide separation. Absorbance was measured at 220 nm and is plotted as millivolts.

26 and 28 were too low in cycles seven and eight to conclude that these sequences terminated after step eight.

The sequences obtained clearly indicate a peptide motif consisting of proline at position 2 and a COOH-terminal hydrophobic residue, predominantly leucine, phenylalanine, or methionine. In addition, a preference for glutamine or asparagine at positions 3 and 8, and a predilection for a second proline were noted. In one case (fraction 24), a small amount of lysine was identified with the tenth degradative step. The H-2L^d motif has also been found in two previously described H-2L^d restricted peptides: that of the MCMV pp89 immediate early regulatory protein, pMCMV (51), and that of

Table 1. *Sequences of Eluted Self Peptides*

Fraction	Sequence										Homology
	1	2	3	4	5	6	7	8	9	10	
21	no sequence										
22	no sequence										
23	X	P	l	E	A	N	Y	Q	X	f	Mouse leukocyte common antigen (6/8)
24	A	P	Q	P	G	M	E	N	F	k	Bovine endozepine (5/10)
25	q	P	Q	R	G	R	E	N	F		*S. erythraeus* rRNA adenine *N*-6-methyltransferase (6/9)
26	X	P	q	P	G	R	E	Q			Mouse proline-rich protein (6/7)
27	X	P	Q	P	n	l	Y	Q	l		Chicken class I B chain precursor (6/8)
28	X	P	a	X	a	y	p	Y			*B. stearothermophilus* superoxide dismutase (5/6)
29	Y	P	N	V	N	I	H	N	F		Mouse Duchenne muscular dystrophy (6/9)
30	X	P	Q	K	A	G	G	F	L	M	Mouse phosphoglycerate kinase (9/9)
											viral antigens
	Y	P	H	F	M	P	T	N	L		pMCMV pp89 168–176
	R	P	Q	A	S	G	V	Y	M		pLCMV NP 118–126

The material in recovered HPLC fractions 21–30 was sequenced by automated Edman degradation. The typeface classifications for amino acids reported include: boldface for a unique residue; plain uppercase for a quantitatively dominant residue; and plain lowercase for a residue with the least assurance because of low quantity or heterogeneity at that position. Positions for which an assignment could not be confidently made are designated with an X. Homologous proteins were identified by searching the NBRF data bank. The homologous proteins are followed by the fraction of exact amino acid matches with identified residues in the sequenced fraction. Accession numbers for best matches were: fraction 23, MUSLCA26 (74); fraction 24, A32944 (75); fraction 25, XYSMRE (76); fraction 26, A29149 (77); fraction 27, SO1172 (78); fraction 28, DSBSNF (79); fraction 29, B27162 (80); and fraction 30, A25567 (81).

the murine lymphocyte choriomeningitis virus, pLCMV (52) (Table 1).

To identify possible sources for the copurified peptides, the NBRF protein data base was scanned for these sequences. The best protein matches and fractions of homology are indicted in Table 1. Only one completely homologous protein, namely phosphoglycerate kinase, a cytosolic protein, was identified. Homologous proteins for the identified sequences should have been readily found if these eluted peptides had originated from serum components.

Binding of Synthetic Peptides to H-2L^d_s. To verify that the eluted peptides were capable of binding to H-2L^d_s, we used a set of synthetic analogues representing the sequenced peptides. When ambiguity at particular sequencing steps did not permit a firm assignment, alanine was substituted in the synthetic peptide to preserve chirality and to impose minimal bias upon the propensity to bind. One of the synthetic peptides, p29 (named for the eluted fraction number), had an identical HPLC retention time as the eluted, sequenced peptide. Others differed, suggesting that the alanine substitutions were not the original residues, or that in some cases the length of the predominant species differed from the apparent consensus sequence length.

The synthetic peptides were analyzed for their ability to bind to purified H-2L^d_s in cold competition with radiolabeled pMCMV (Fig. 2, A–C). Four of the synthetic peptide

analogues, three nonamers and one decamer (p25A, p27, p29, and p30; Fig. 2 A), competed better than cold pMCMV in solution binding to H-2L^d_s, and two of the decamer peptides (p23 and p24A; Fig. 2, A and B) did not compete. One of the two decamer peptides that did not compete (p24A) was resynthesized as a nonamer truncation, p24B, which competed effectively (Fig. 2 B). Thus the failure of peptide p24A to bind was due to the lysine residue at its COOH terminus. An H-2L^d core motif peptide, APAAAAAAL (abbreviated pAPAL) competed for binding within an order of magnitude of concentration of pMCMV.

A second and independent assay of peptide binding to soluble H-2L^d_s was employed, measuring the ability of the synthetic peptide to induce the formation of the conformational epitope bound by the mAb 30–5–7S (49), using real time surface plasmon resonance (see Materials and Methods) (Fig. 2, D–F) (27–31). Normalized RUs provided a quantitative indication of the mass of H-2L^d_s associated with the antibody at a given time. The data in Fig. 2 D show that p27, p29, and p30 were slightly more effective in inducing the formation of the 30–5–7S epitope than pMCMV, whereas another longer H-2L^d restricted peptide, ptum 91A- (53), minimally induced the epitope above the basal level. In Fig. 2 E, we observe that p24B and p25B were equivalent to pMCMV, whereas p24A (the decamer with the COOH-terminal lysine) was substantially worse (the half-maximal

Table 2. *Calculated Contacts and Surface Area of MHC-bound Peptides*

| Peptide | Sequence | Atomic contacts | | | Surface | | Relative binding | |
		van der Waals*	H bonds‡	Ion pairs§	Buried (Å²)‖	Exposed (Å²)‖	Competition¶	Epitope induction**
p25A	QPQRGRENF	158	18	2	587	418	+ + +	+
p27	APQPNLYQL	132	10	1	498	445	+ + + + (3.9)	+ +
p29	YPNYNIHNF	199	10	1	589	438	+ + + (1.8)	+ +
pMCMV	YPHFMPTNL	196	15	1	608	375	+ + (1.0)	+ +
p24B	APQPGMENF	120	10	1	524	372	+ (0.4)	+ +
pLCMV	RPQASGVYM	145	15	2	590	324	ND	+
p25B	APQRGRENF	125	13	1	535	403	ND	+ +
p30	APQKAGGFLM	138	14	1	572	387	+ + + +	+ +
pHLAB27	ARYAASTEL	125	13	1	556	371	ND	ND

* Two atoms are said to be in van der Waals contact if they are within the sum of their van der Waals' radii + 0.5 Å of each other.
‡ Two atoms are said to be hydrogen-bonded if they are within 2.9 ± 0.5 Å of each other.
§ Two oppositely charged atoms are said to form an ion pair if they are within 2.85 ± 0.5 Å of each other.
‖ Molecular surface available to a spherical probe of 1.7 Å radius was calculated for the indicated peptide/MHC models and free peptides using the Connolly algorithm (82) in QUANTA 3.2. The exposed surface area of peptide in the complex is listed. The difference between this area and total surface area of the free peptide is tabulated as buried surface area.
¶ Relative binding was assessed by competition for solution binding of radiolabeled pMCMV for H-2L^d_s using spin columns as described in Fig. 2. For those peptides for which competition was assessed over a range of three orders of magnitude, the value of the K_d for a single site fit was compared with that of pMCMV as a reference and is given in parenthesis. p25A and p30 were assigned the semi-quantitative values based on graphic comparison of the other peptides. Binding was quantitatively compared using both ALLFIT (83) and LIGAND (25). For these competitive curves, the data were insufficient for statistically significant discrimination between one- and two-site binding models. Binding is listed relative to the competition with pMCMV: +, 0.1–0.5-fold the binding with pMCMV; + +, 0.6–1.5; + + +, 1.6–2.0; and + + + +, >2.1-fold pMCMV binding.
** Relative binding was assessed by induction of the 30-5-7S epitope as described in Fig. 2. Half-maximal values for epitope induction are listed as: +, <50% the binding in the pMCMV; or + +, >50% the binding with pMCMV.
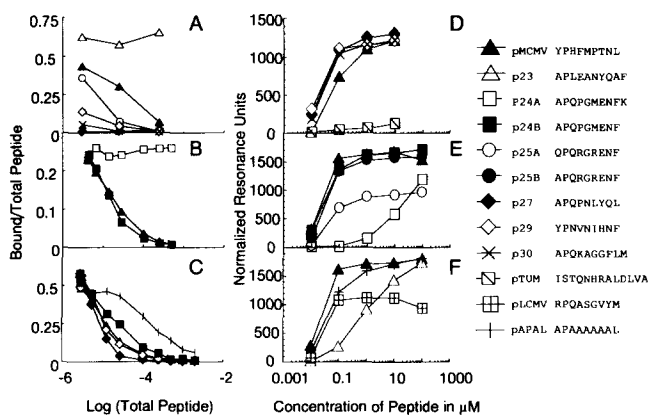
Figure 2. Synthetic analogues of eluted self peptides bind to H-2L^d_s molecules. Synthetic peptides as indicated were used to compete for binding of ^{125}I-labeled pMCMV to H-2L^d_s in solution (A–C), as described in Materials and Methods. The results of three separate experiments are shown in A–C, respectively. pMCMV (▲), as a cold competitor, was the standard in each of the three experiments. Names and sequences of other peptides are indicated. Each data point represents the mean of triplicate determinations. Peptides were also used to induce the α_2 domain conformationally sensitive 30-5-7S epitope on H-2L^d_s molecules as measured by real time surface plasmon resonance (D–F; see Materials and Methods). Varying concentrations of peptide were added in solution to purified H-2L^d_s molecules. Normalized RUs represent the difference in RUs in the presence and absence of added peptide.

concentration of peptide required for epitope induction is about three orders of magnitude higher). pLCMV and p25A exhibited a peculiar behavior in that they induced the formation of the 30-5-7S epitope at low concentrations, but saturated at a lower plateau (Fig. 2, E and F). As shown in Fig. 2 F, pAPAL was detectably and p23 substantially less effective in inducing the epitope than pMCMV.

To evaluate the binding behavior of these synthetic peptides in a functionally relevant manner, and to confirm that peptides eluted from a soluble analogue of H-2L^d could bind cell surface H-2L^d, we used them to inhibit the presentation of the pMCMV peptide by an H-2L^d-expressing L cell transfectant, T1.1.1, to a pMCMV-specific, H-2L^d-restricted T cell hybridoma, E1B6 (21). Four of the six synthetic peptides tested (p25A, p27, p29, and p30) inhibited the presentation of pMCMV to the hybridoma, whereas two peptides (p23 and p24A), which had failed to bind in either the physical or the serological assays, were also unable to compete functionally (data not shown).

Sequence and Structural Variability of H-2L^d, Binding Peptides. With the demonstration of the binding of several viral and self peptides, it is valuable to analyze the location and degree of variability of their amino acids. Variability based on seven nonamer peptides shown to bind was calculated based on two algorithms: sequence variability was computed according to Wu and Kabat (32), and structural variability was computed according to Padlan (33) (Fig. 3). Wu-Kabat variability provides an indication of how many different amino acids are observed at a given position relative to the frequency of the most common amino acid at that position. This calculation treats both conservative and radical substitutions equivalently. Structural variability as calculated according to

Padlan imposes a bias dependent on the degree of chemical dissimilarity of the observed substitutions. Peptide amino acids at position 2 showed no variability by either measure, and positions 3, 8, and 9 had limited variability. In contrast, the remaining positions exhibited substantial variability, with positions 1, 4, and 6 significantly greater by both algorithms.

Molecular Modeling of the Peptide/MHC Complex. In the absence of a high resolution x-ray crystallographic structure for H-2L^d, we have built molecular models of the H-2L^d_s/β_2-m/peptide complex. Madden et al. (14) have published the x-ray structure of the human class I MHC molecule HLA-B27 complexed to human β_2-m, and have modeled the electron density in the binding cleft to a nonamer peptide, ARYAASTEL. Since the overall sequence homology of H-2L^d_s to HLA-B27 over the 275 residues of the crystallized H chain is 76% with conservation of features such as the location of cysteine residues involved in intrachain disulfide bonds, it is reasonable to employ a mixture of heuristic and objective methods to develop such a model (54–57). Beginning with the coordinates derived from 3.0 Å resolution x-ray diffraction data of the HLA-B27/hβ_2-m/peptide complex, provided by D. Madden and D. Wiley, the H-2L^d_s/mβ_2-m/peptide models were built by substituting the polymorphic amino acid residues of H-2L^d_s, mβ_2-m, and each of a set of H-2L^d restricted peptides, followed by regularization, energy minimization, and simulated annealing (see Materials and Methods). Independently derived models of the H-2L^d_s/mβ_2-m/peptide complexes were analyzed for atomic contacts between the peptide and the H chain, and for buried and exposed surface area of the peptides (see Table 2). The number and kind of such contacts and the area of buried or exposed surface did not clearly correlate with the relative binding of the peptides to the class I molecule.

The seven nonamer peptides shown to bind to H-2L^d_s were also analyzed for the contributions of each residue to deviation from average position, average fractional accessibility to solvent, and total number of atomic contacts, and these were compared with the sequence and structural variabilities calculated above (see Table 3). In general, variability paralleled the rms deviation from the average position, with the exception of position 1, where variability was permitted,
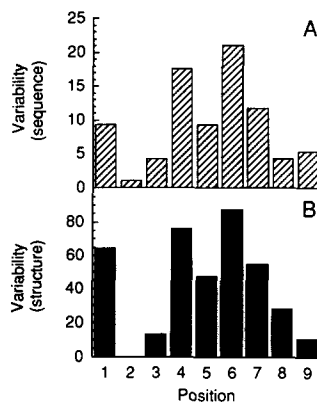

Figure 3. Sequence and structural variability of seven H-2L^d-binding nonamer peptides. Sequences of peptides p25A, p27, p29, pMCMV, p24B, pLCMV, and p25B were compared according to the methods of Wu and Kabat (32) (A) and Padlan (33) (B).

Table 3. *Parameters of the Seven H-2L^d-binding Peptides*

Position	Sequence variability*	Structural variability‡	rms deviation from average position (Å)§	Average fractional accessibility (±SD)‖	Total no. atomic contacts¶
1	9.3	64.6	0.61	0.09 ± 0.10	275
2	1.0	0.0	0.79	0.00 ± 0.00	109
3	4.2	13.0	0.91	0.24 ± 0.12	128
4	17.5	75.9	0.93	0.91 ± 0.10	40
5	9.3	47.2	0.85	0.67 ± 0.15	46
6	21.0	87.3	1.06	0.90 ± 0.09	53
7	11.7	54.9	1.06	0.45 ± 0.16	115
8	4.2	27.8	0.67	0.74 ± 0.06	37
9	5.3	10.7	0.62	0.07 ± 0.03	272

Molecular models of each of the seven nonamer peptides, p25A, p27, p29, pMCMV, p24B, pLCMV, and p25B, bound to H-2L^d, and mβ_2-m were constructed as described in Materials and Methods. Indicated parameters were calculated as follows:
* Sequence variability was computed using the formula of Wu and Kabat (32).
‡ Structural variability quantifies the structural differences among the various amino acid residues occurring at each position and was computed according to Padlan (33) using the amino acid dissimilarity values of Grantham (34).
§ Deviation of C$_\alpha$ position from the average was calculated.
‖ The fractional accessibility of each amino acid residue, defined as the amount of surface area that is accessible to a solvent probe in the three-dimensional structure divided by the surface are that would be exposed if the residue (X) were in an isolated Gly-X-Gly tripeptide with the same backbone configuration as the corresponding tripeptide in the structure, was computed according to Padlan (84).
¶ Two atoms are said to be in contact if they are within the sum of their van der Waals' radii + 0.5Å of each other.

but large deviation in position was not observed. Residues 1, 2, and 9 have very little exposed surface area on average, and residues 3–8 are generally exposed. The number of atomic contacts, as expected, inversely correlated with the exposed surface area: the most contacts are with atoms from residues 1, 2, 3, 7, and 9.

Illustrations of the computer models of two peptide/H-2L^d, complexes, the pMCMV nonamer and the p30 decamer, are shown in Fig. 4. In each of these panels, the stick figure of peptide including its side chains is shown with a partial α carbon tracing of the H-2L^d, molecule. In Fig. 4, A and C, the side chain of the invariant residue, tryptophan 147, is shown for reference. In addition, the solvent-exposed surface, as calculated by the Connolly algorithm (48), of the peptide residues in the complex is shown. In the pMCMV/H-2L^d, model (Fig. 4, A and B), the side chain of the first residue of the peptide (tyrosine) is partially accessible to solvent, and the side chains of residues 4, 5, 6, and 7 are largely accessible to solvent. The model of the p30 decamer, which has a potentially flexible glycine-glycine sequence at positions 6 and 7, is also well accommodated in the binding cleft (Fig.

4, C and D). For comparison, the peptides of the pMCMV and p30 complexes have been superimposed (Fig. 5) to illustrate the similarity of the backbone trace despite the difference in length of one amino acid.

Discussion

In this paper we report the isolation of self peptides from a soluble analogue of the murine MHC class I molecule, H-2L^d_s, their sequence determination, and the rebinding of synthetic peptides corresponding to those sequences to both purified soluble and cell surface H-2L^d. These peptides included both nonamers and decamers, consistent with the findings of others who purified MHC class I from detergent extracts of cells and analyzed the peptides eluted from such preparations (2, 4, 58). Unlike viral peptides eluted from H-2K^b (5) no octamers were clearly identified.

A number of laboratories have examined self and viral peptides for their length and primary sequence, but few have reevaluated the sequenced peptides for the ability to bind to the MHC molecule from which they were eluted. As the

Figure 4. Molecular graphics models of the pMCMV/H-2L^d, and p30/H-2L^d, complexes. Computer models of pMCMV (YPHFMPTNL) complexed to H-2L^d, (A and B), and p30 (APQKAGGFLM) complexed to H-2L^d, (C and D) were generated as described in Materials and Methods. α carbon backbone structure of the α1 and α2 helices and the invariant tryptophan side chain 147 are indicated in the en face view (A and C). The side views (B and D) show the β-sheet floor and only the α1 domain for clarity. (Dots) Solvent-accessible surfaces of the bound peptides.

Figure 5. Structure of the modeled pMCMV (nonamer) and p30 (decamer) peptides. Computer models were generated as described in Materials and Methods. The bound peptides of the modeled complexes are superimposed in this display. pMCMV (yellow); p30 (red). The NH₂ termini (tyrosine for pMCMV, and alanine for p30) are to the left, and the COOH termini are to the right.

A

α1

TYR-1   LEU-8

α2

B

TYR-1   LEU-8

C

α1

ALA-1   MET-10

α2

D

MET-10

ALA-1
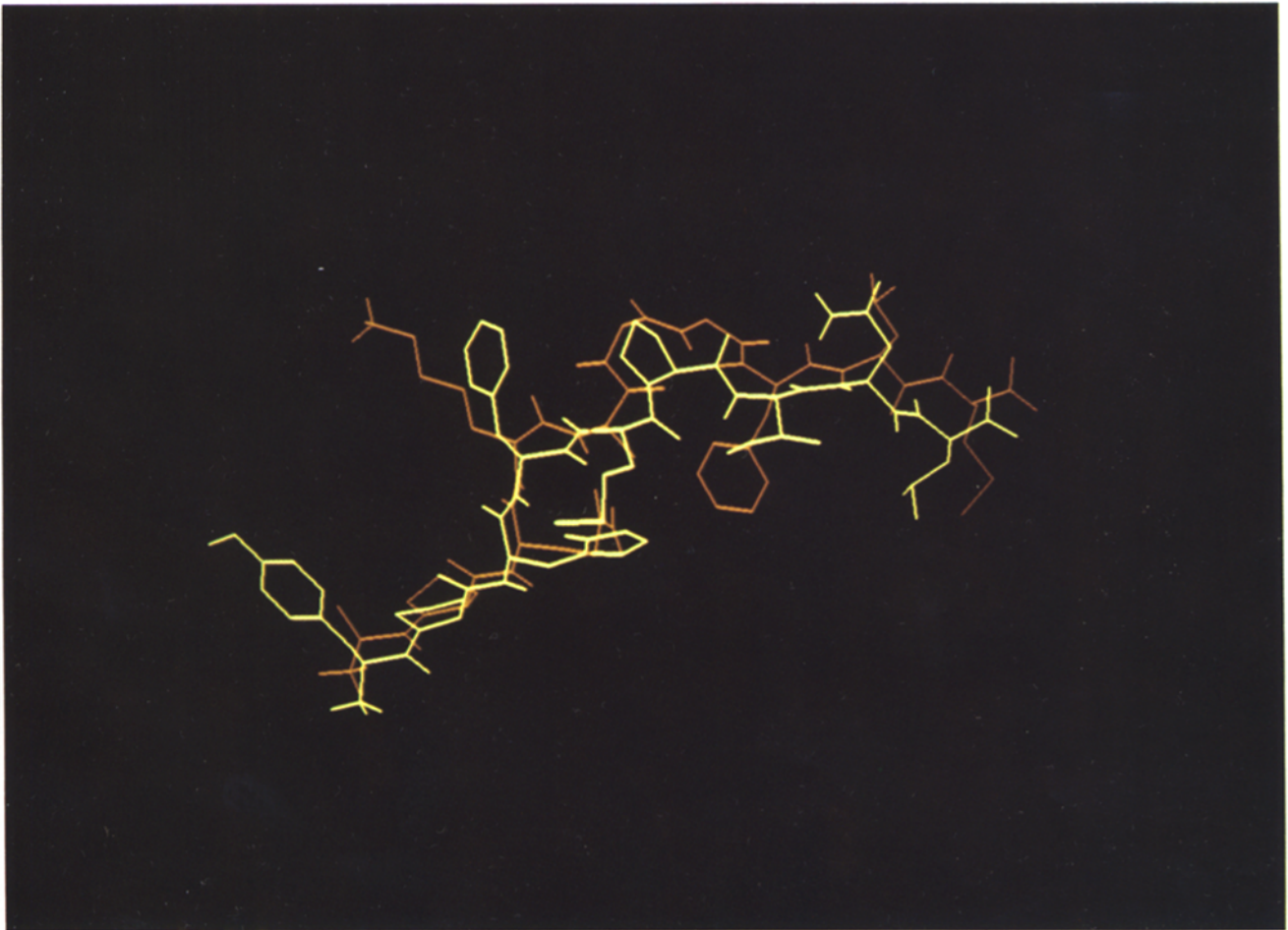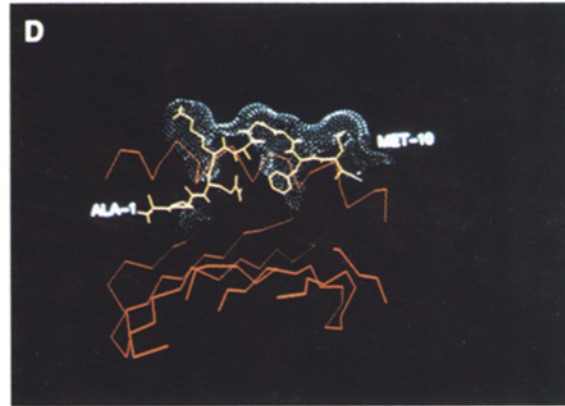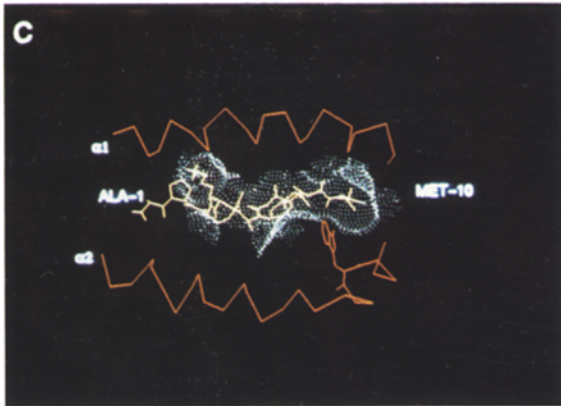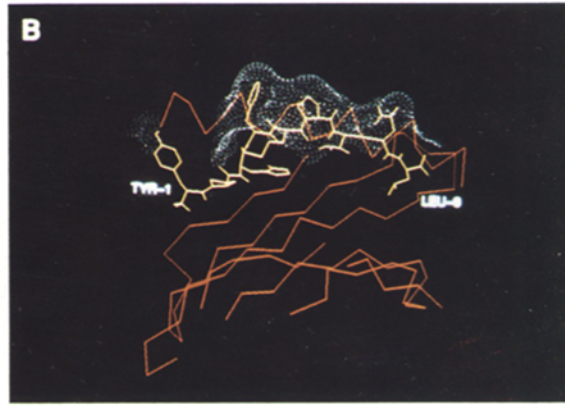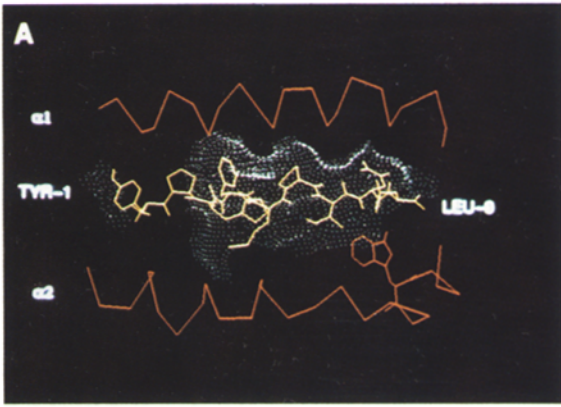
quantitation of such eluted peptides is difficult and the yield of sequential Edman degradative steps may appear to indicate termination prematurely, a second step to verify that these peptide sequences represent bound peptides was necessary. We have demonstrated binding of synthetic analogues of the H-2L^d_s-eluted peptides using several independent assays: solution binding by competition, epitope induction using surface plasmon resonance to evaluate the presence of the α_2 domain 30–5–7S conformationally sensitive epitope, and competition for presentation by an H-2L^d-expressing L cell. Overall, these peptides bound at apparent affinities comparable with those of the reference peptide, pMCMV.

Binding in solution permitted quantitative comparison of the binding of the test peptides as compared with the reference ^{125}I-labeled peptide, pMCMV. Several peptides bound better than pMCMV: p25A, p27, p29, and p30. Competition by the decamer peptide p30 supported the view that longer peptides may bind. However, the inability of the other two decamer peptides, p23 and p24A, to compete with pMCMV for binding illustrated the potential difficulties in identifying peptides by protein sequencing of HPLC fractions of acid-eluted material. The decamer p24A, with a questionable assignment of lysine at the COOH terminus, did not bind; however, the nonamer p24B terminating with phenylalanine did. The identification of the permissive COOH terminus of this peptide supported the dominant motif for H-2L^d binding, namely a proline at position 2 and a leucine, phenylalanine, or methionine at position 9 or 10. In addition, the motif peptide, pAPAL, bound well in the solution assay, although about an order of magnitude worse than pMCMV.

The results of the solution binding assay were qualitatively consistent with those of the epitope induction assay detecting the conformational α_2 domain epitope, 30–5–7S. In the solution binding assay (Fig. 2, A–C) p23 and p24A did not compete detectably for binding by pMCMV, but in the epitope induction assay (Fig. 2, D–F), both peptides (at about 100- and 1,000-fold the pMCMV concentration) could induce the epitope. Whether the induction of epitope directly and quantitatively correlates with peptide binding is not yet established, and the apparent quantitative discrepancies between the solution binding assay and the epitope induction assay may indicate that the two are not firmly linked.

The motif was also confirmed by the demonstration that a motif peptide, pAPAL (APAAAAAAL), could effectively bind both in the solution binding and epitope induction assays. This finding is consistent with that of Maryanski et al. (59) who demonstrated in functional competition assays that H-2K^d could bind the consensus peptide AYPPPPPTLA or AYGGGGGGTLA, and that of Jardetzky et al. (60) who demonstrated the binding of a similar substituted motif peptide (AAYAAAAAAKAAA) to the class II MHC molecule, HLA-DR1.

To examine the implied structural constraints of the eluted self peptides, we built eight independent computer models of the H-2L^d_s/mβ_2-m/peptide complex. Our models, which take advantage of the similarity of the primary structures of the mouse H-2L^d_s and human HLA-B27 MHC class I molecules, and the HLA-B27 structure, depict the bound pep-

tides in extended conformations. These models contrast markedly with one recently presented by Rognan et al. (61) in which the pMCMV/H-2L^d complex was modeled from first principles. In that model, the peptide was built as an α-helix without benefit of the crystallographic data of MHC bound peptides. Although considerable evidence implicates the propensity for a helical or amphipathic helical structure as a predictor of MHC-restricted antigenic peptides (62–65), the analysis of functional MHC-restricted peptide binding (59, 66, 67), of the solution structure of antigenic peptides thought to be helical by predictive algorithms (68), and of the three-dimensional structure of peptides bound to HLA class I molecules (12, 14), and more recently H-2K^b (15), indicates that the peptide is probably in an extended conformation.

The apparent requirement for proline at position 2 raises the possibility of at least two contributing factors: (a) the specific geometry, particularly with respect to peptide residue 1, contributed by the proline in either a cis or trans configuration; and (b) the spatial conformation and charge of the region of H-2L^d that putatively interacts with the side chain of position 2. Our molecular model of pMCMV/H-2L^d_s indicates that the NH$_2$-terminal amine of the peptide is hydrogen bonded to the hydroxyl oxygen of tyrosines 7, 59, and 171. The geometry of proline at 2 constrains the position of the tyrosine at position 1 to allow these bonds. The N-C$_\alpha$-C dihedral angle of the arginine at peptide position 2 of HLA-B27 has a similarly restrained orientation as the trans proline in our models of the H-2L^d_s-bound peptides. It is interesting that the energy stabilization contributed by the hydrogen bonding of the NH$_2$-terminal amine is consistent with the findings of Fersht et al. (69) who found a difference of $\sim$3 kcal mol^{-1} for the energy contributed by a hydrogen bond to a charged, as compared to an uncharged donor or acceptor. Thus, a hydrogen bond to a charged NH$_2$-terminal amine is about 3 kcal mol^{-1} stronger than one to an uncharged amine such as one in a peptide bond.

The sequence restrictions at positions 2 and 9 found in our motif can be further understood by inspecting the H-2L^d_s/peptide models for the areas of H chain where these peptide residues interact, namely the B and F pockets. In the H-2L^d_s/pMCMV model, atoms of the H chain that are in contact with the proline derive from tyrosine 7, arginine 62, isoleucine 63, isoleucine 66, tyrosine 99, tyrosine 159, and glutamic acid 163. In contrast with the B pocket of HLA-B27 which is lined by the following residues: tyrosine 7, histidine 9, threonine 24, valine 25, valine 34, glutamine acid 45, glutamic acid 63, isoleucine 66, cysteine 67, lysine 70, and tyrosine 99, the corresponding region of H-2L^d_s differs with glutamic acid at 9, serine at 24, tyrosine at 45, isoleucine at 63, alanine at 67, and glutamine at 70. Thus, the B pocket of H-2L^d is partially closed off by the bulky tyrosine side chain at position 45. This may restrict the size of the side chain that can be easily accommodated by the pocket. In addition, the salt bridge linking the α_1 to the α_2 helix, arginine 62 to glutamic acid 163, provides a canopy that constrains the proline at position 2. It is likely that for some peptides, serine or alanine at position 2 is acceptable, since

alanine and serine have side chain volumes and hydrophobicity very similar to proline (34, 70). Evidence that serine is permitted at position 2 includes the identification of a self peptide LSPFPFDL by Udaka et al. (71) that sensitizes H-2L^d-expressing cells for recognition by an allospecific CTL clone. The tum 91A$^-$ peptide, ISTQNHRALDLVA, described by Lurquin et al. (53), and shown to induce the 30-5-7S epitope on H-2L^d positive cell surfaces in the presence of serum (24) has a serine in position 2, and most likely, the minimal core peptide is ISTQNHRAL (Corr, M., and D. Margulies, unpublished data). In addition, a pMCMV analogue with alanine at position 2, although it has drastically decreased activity, still binds detectably (72) (Boyd, L., M. Corr, S. Kozlowski, R. Ribaudo, and D. Margulies, unpublished data).

The identity of the COOH-terminal residue as leucine, phenylalanine, or methionine is consistent with the findings of Bordo and Argos (73) who demonstrated the constraints on allowed amino acid substitutions in proteins. They observed the likelihood of substitution of isoleucine, leucine, valine, and methionine for one another, and noted the ability of phenylalanine to substitute for leucine in protein cores. Scrutiny of the region of H-2L^d_s, which interacts with the COOH-terminal residue (the F pocket of Garrett et al. [13]) reveals that the hydrophobic side chain of the peptide protrudes into a pocket lined by tryptophan 73, leucine 81, phenylalanine 116, tyrosine 123, threonine 143, and tryptophan 147 (which hydrogen bonds to the carbonyl oxygen of the eighth residue of the peptide). Leucine fits well into this pocket, and the model of p29 with H-2L^d_s indicates that the COOH-terminal phenylalanine of p29 can also stack with tryptophan 73, phenylalanine 116, and tryptophan 147.

Computer graphic depiction of the peptide binding groove with the pMCMV peptide in place is shown in Fig. 4. In these views, the H-2L^d_s α-carbon backbone and the side chain of the invariant tryptophan at position 147 (Fig. 4 A only) are shown. Solvent accessibility is depicted by the dot surface. The side chain of the NH$_2$-terminal tyrosine residue at position 1 is accessible to solvent, whereas the NH$_2$-terminal alanine of the decamer is not. The decamer model, in which the NH$_2$ and COOH termini were first superimposed on the first and ninth residues of the HLA-B27 bound peptide, and the intervening residues were then placed, reveals similarity in the basic backbone trace, likely due to the

flexibility of the glycine-glycine sequence at positions 6 and 7 (Fig. 5). Thus, reasonable models for decamer peptide/class I complexes can be generated by fixing the appropriate peptide residues in the B and F pockets first.

The comparison of the sequences and the structural models was carried out in two ways. First, we analyzed the peptides in toto and compared each model with the others (Table 2). No striking predictor of strength of binding was observed. When the individual peptides were compared with one another on a residue by residue basis, however (Table 3), it became clear that residues that permit sequence and structural variability also have greater deviation from the mean position, greater fractional exposed surface area, and fewer atomic contacts. Conversely, residues with lower variability in general make more atomic contacts with the MHC molecule, are spatially more constrained, and are less exposed to solvent.

While this manuscript was under review, the structures of two complexes of viral peptides with the mouse H-2K^b molecule were published (15, 16). These structures and those that we have modeled show remarkable similarity. In particular, the backbones of the peptides are in an extended conformation, and prolines are in *trans*. Longer peptides are accommodated because of their central flexibility.

In conclusion, we have demonstrated that an immunoadsorbent purified, soluble analogue of the murine MHC class I molecule, H-2L^d_s, contains bound self peptides that can be eluted with acid, fractionated by HPLC, and sequenced. Synthetic peptides corresponding to these sequences can be tested for binding by any of several methods, and three-dimensional models of the peptide/MHC complexes can be made by substitution of residues onto the HLA-B27/peptide crystallographic coordinates, followed by simulated annealing and energy minimization. The resulting models permitted comparison of sequence and structural variability with solvent accessibility of the bound peptide and spatial deviation of specific amino acid residues. Such analysis permitted a unified view of the rules that govern MHC/peptide binding: specific interactions in pockets as proposed by Garret et al. (13), and variability in the nonpocket regions. Such determination of constrained and variable portions of MHC-restricted antigenic peptides, in addition to furthering our overall understanding of the way these unique peptide receptors function, should permit the design of peptides for blocking or enhancing immune responses.

Address correspondence to Dr. David H. Margulies, Bldg. 10, Rm. 11N311, NIH, Bethesda, MD 20892.

References

1. Bjorkman, P.J., and P. Parham. 1990. Structure, function, and diversity of class I major histocompatibility complex molecules. *Annu. Rev. Biochem.* 59:253.

2. Falk, K., O. Rötzschke, 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.

3. 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.

4. Hunt, D.F., R.A. Henderson, J. Shabanowitz, K. Sakaguchi, H. Michel, N. Sevilir, A.L. Cox, E. Appella, and V.H. Engelhard. 1992. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science (Wash. DC).* 255:1261.

5. Van Bleek, G.M., and S.G. Nathenson. 1990. Isolation of an endogenously processed immunodominant viral peptide from the class I H-2K^b molecule. *Nature (Lond.).* 348:213.

6. Falk, K., O. Rötzschke, K. Deres, J. Metzger, G. Jung, and H.-G. Rammensee. 1991. Identification of naturally processed viral nonapeptides allows their quantification in infected cells and suggests an allele-specific T cell epitope forecast. *J. Exp. Med.* 174:425.

7. Pamer, E.G., J.T. Harty, and M.J. Bevan. 1991. Precise prediction of a dominant class I MHC-restricted epitope of *Listeria monocytogenes. Nature (Lond.).* 353:852.

8. Schwartz, R.H. 1985. T-lymphocyte recognition of antigen in association with gene products of the major histocompatibility complex. *Annu. Rev. Immunol.* 3:237.

9. Townsend, A., and H. Bodmer. 1989. Antigen recognition by class I-restricted T lymphocytes. *Annu. Rev. Immunol.* 7:601.

10. Bjorkman, P.J., M.A. Saper, B. Samraoui, W.S. Bennett, J.L. Strominger, and D.C. Wiley. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature (Lond.).* 329:506.

11. Bjorkman, P.J., M.A. Saper, B. Samraoui, W.S. Bennett, J.L. Strominger, and D.C. Wiley. 1987. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature (Lond.).* 329:512.

12. Saper, M.A., P.J. Bjorkman, and D.C. Wiley. 1991. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J. Mol. Biol.* 219:277.

13. Garrett, T.P., M.A. Saper, P.J. Bjorkman, J.L. Strominger, and D.C. Wiley. 1989. Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature (Lond.).* 342:692.

14. Madden, D.R., J.C. Gorga, J.L. Strominger, and D.C. Wiley. 1991. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature (Lond.).* 353:321.

15. Fremont, D.H., M. Matsumura, E.A. Stura, P.A. Peterson, and I.A. Wilson. 1992. Crystal structures of two viral peptides in complex with murine MHC class I H-2K^b. *Science (Wash. DC).* 257:919.

16. Matsumura, M., D.H. Fremont, P.A. Peterson, and I.A. Wilson. 1992. Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science (Wash. DC).* 257:927.

17. Rötzschke, O., K. Falk, K. Deres, H. Schild, M. Norda, J. Metzger, G. Jung, and H.-G. Rammensee. 1990. Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells [see comments]. *Nature (Lond.).* 348:252.

18. Hunt, D.F., M. Hanspeter, T.A. Dickinson, J. Shabanowitz, A.L. Cox, K. Sakaguchi, E. Appella, H.M. Grey, and A. Sette.

1992. Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad. *Science (Wash. DC).* 256:1817.

19. Margulies, D.H., L.F. Boyd, S. Kozlowski, L. Kjer-Nielsen, R. Lopez, J. Schneck, and R. Hunziker. 1990. Multivalent requirement for the stimulation of alloreactive T cells: studies with engineered soluble MHC class I proteins in vitro and in vivo. *In* Transgenic Mice and Mutants in MHC Research. I.K. Egorov and C.S. David, editors. Springer-Verlag, Berlin. 39–46.

20. Ozato, K., T.H. Hansen, and D.H. Sachs. 1980. Monoclonal antibodies to mouse MHC antigens. II. Antibodies to the H-2L^d antigen, the products of a third polymorphic locus of the mouse major histocompatibility complex. *J. Immunol.* 125:2473.

21. Boyd, L.F., S. Kozlowski, and D.H. Margulies. 1992. Solution binding of an antigenic peptide to a major histocompatibility complex class I molecule and the role of β_2-microglobulin. *Proc. Natl. Acad. Sci. USA.* 89:2242.

22. Evans, G.A., D.H. Margulies, B. Shykind, J.G. Seidman, and K. Ozato. 1982. Exon shuffling: mapping polymorphic determinants on hybrid mouse transplantation antigens. *Nature (Lond.).* 300:755.

23. McCluskey, J., L. Boyd, M. Foo, J. Forman, D.H. Margulies, and J.A. Bluestone. 1986. Analysis of hybrid H-2D and L antigens with reciprocally mismatched aminoterminal domains: functional T cell recognition requires preservation of fine structural determinants. *J. Immunol.* 137:3881.

24. Lie, W.-R., N.B. Myers, J. Gorka, R.J. Rubocki, J.M. Connolly, and T.H. Hansen. 1990. Peptide ligand-induced conformation and surface expression of the L^d class I MHC molecule. *Nature (Lond.).* 344:439.

25. Munson, P.J., and D. Rodbard. 1980. Ligand: a versatile computerized approach for characterization of ligand-binding systems. *Anal. Biochem.* 107:220.

26. DeLean, A., P.J. Munson, and D. Rodbard. 1978. Simultaneous analysis of families of sigmoidal curves: application to bioassay, radioligand assay, and physiological dose–response curves. *Am. J. Physiol.* 235:E97.

27. Mayo, C.S., and R.B. Hallock. 1989. Immunoassay based on surface plasmon oscillations. *J. Immunol. Methods.* 120:105.

28. Fägerstam, L.G., A. Frostell, R. Karlsson, M. Kullman, A. Larsson, M. Malmqvist, and H. Butt. 1990. Detection of antigen–antibody interactions by surface plasmon resonance. Application to epitope mapping. *J. Mol. Recognition.* 3:208.

29. Pollard-Knight, D., E. Hawkins, D. Yeung, D.P. Pashby, M. Simpson, A. McDougall, P. Buckle, and S.A. Charles. 1990. Immunoassays and nucleic acid detection with a biosensor based on surface plasmon resonance. *Ann. Biol. Clin.* 48:642.

30. Jönsson, U., L. Fägerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Löfås, B. Pearson, H. Roos, I. Rönnberg, et al. 1991. Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. *Biotechniques.* 11:620.

31. Johnsson, B., S. Löfås, and G. Lindquist. 1991. Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal. Biochem.* 198:268.

32. Wu, T.T., and E.A. Kabat. 1970. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* 132:211.

33. Padlan, E.A. 1977. Structural implications of sequence vari-

ability in immunoglobulins. *Proc. Natl. Acad. Sci. USA.* 74:2551.

34. Grantham, R. 1974. Amino acid difference formula to help explain protein evolution. *Science (Wash. DC).* 185:862.

35. Evans, G.A., D.H. Margulies, R.D. Camerini-Otero, K. Ozato, and J.G. Seidman. 1982. Structure and expression of a mouse major histocompatibility antigen gene, H-2L^d. *Proc. Natl. Acad. Sci. USA.* 79:1994.

36. Moore, K.W., B.T. Sher, Y.H. Sun, K.A. Eakle, and L. Hood. 1982. DNA sequence of a gene encoding a BALB/c mouse Ld transplantation antigen. *Science (Wash. DC).* 215:679.

37. Watts, S., C. Wheeler, R. Morse, and R.S. Goodenow. 1989. Amino acid comparison of the class I antigens of mouse major histocompatibility complex. *Immunogenetics.* 30:390.

38. Pullen, J.K., R.M. Horton, Z.L. Cai, and L.R. Pease. 1992. Structural diversity of the classical H-2 genes: K, D, and L. *J. Immunol.* 148:953.

39. Mellor, A.L., E.H. Weiss, M. Kress, G. Jay, and R.A. Flavell. 1984. A nonpolymorphic class I gene in the murine major histocompatibility complex. *Cell.* 36:139.

40. Gates, F.T.D., J.E. Coligan, and T.J. Kindt. 1981. Complete amino acid sequence of murine β_2-microglobulin: structural evidence for strain-related polymorphism. *Proc. Natl. Acad. Sci. USA.* 78:554.

41. Parnes, J.R., and J.G. Seidman. 1982. Structure of wild-type and mutant mouse β_2-microglobulin genes. *Cell.* 29:661.

42. Jones, T.A. 1978. A graphics model building and refinement system for macromolecules. *J. Appl. Crystallography.* 11:268.

43. Powell, M.J.D. 1977. Restart procedures for the conjugate gradient method. *Math. Programm.* 12:241.

44. Brünger, A.T. 1990. X-PLOR Manual, version 2.1.

45. Kirkpatrick, S., C.D.J. Gelatt, and M.P. Vecchi. 1983. Optimization by simulated annealing. *Science (Wash. DC).* 220:671.

46. Wilson, S.R., and W.L. Cui. 1990. Applications of simulated annealing to peptides. *Biopolymers.* 29:225.

47. Brooks, B.R., R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Computational Chemistry.* 4:187.

48. Connolly, M.L. 1983. Analytical molecular surface calculation. *J. Appl. Crystallog.* 16:548.

49. Lie, W.R., N.B. Myers, J.M. Connolly, J. Gorka, D.R. Lee, and T.H. Hansen. 1991. The specific binding of peptide ligand to L^d class I major histocompatibility complex molecules determines their antigenic structure. *J. Exp. Med.* 173:449.

50. Ribaudo, R.K., and D.H. Margulies. 1992. Independent and synergistic effects of disulfide bond formation, β_2-microglobulin, and peptides on class I MHC folding and assembly in an *in vitro* translation system. *J. Immunol.* In press.

51. Reddehase, M.J., J.B. Rothbard, and U.H. Koszinowski. 1989. A pentapeptide as minimal antigenic determinant for MHC class I-restricted T lymphocytes. *Nature (Lond.).* 337:651.

52. Schulz, M., P. Aichele, R. Schneider, T.H. Hansen, R.M. Zinkernagel, and H. Hengartner. 1991. Major histocompatibility complex binding and T cell recognition of a viral nonapeptide containing a minimal tetrapeptide. *Eur. J. Immunol.* 21:1181.

53. Lurquin, C., A. Van Pel, B. Mariamé, E. De Plaen, J.P. Szikora, C. Janssens, M.J. Reddehase, J. Lejeune, and T. Boon. 1989. Structure of the gene of tum- transplantation antigen P91A: the mutated exon encodes a peptide recognized with L^d by cytolytic T cells. *Cell.* 58:293.

54. Browne, W.J., A.C.T. North, D.C. Phillips, K. Brew, T.C. Vanaman, and R.L. Hill. 1969. A possible three-dimensional structure of bovine a-lactalbumin based on that of hen's egg-white lysozyme. *J. Mol. Biol.* 42:65.

55. Padlan, E.A., D.R. Davies, I. Pecht, D. Givol, and C. Wright. 1976. Model-building studies of antigen-binding sites: the hapten-binding site of MOPC-315. *Cold Spring Harb. Symp. Quant. Biol.* 41:627.

56. Blundell, T.L., G. Elliott, S.P. Gardner, T. Hubbard, S. Islam, M. Johnson, D. Mantafounis, P. Murray-Rust, J. Overington, J.E. Pitts, et al. 1989. Protein engineering and design. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 324:446.

57. Greer, J. 1990. Comparative modeling methods: application to the family of the mammalian serine proteases. *Proteins Struct. Funct. Genet.* 7:317.

58. Wei, M.L., and P. Cresswell. 1992. HLA-A2 molecules in an antigen-processing mutant cell contain signal sequence-derived peptides. *Nature (Lond.).* 356:443.

59. Maryanski, J.L., A.S. Verdini, P.C. Weber, F.R. Salemme, and G. Corradin. 1990. Competitor analogs for defined T cell antigens: peptides incorporating a putative binding profile and polyproline or polyglycine spacers. *Cell.* 60:63.

60. Jardetzky, T.S., J.C. Gorga, R. Busch, J. Rothbard, J.L. Strominger, and D.C. Wiley. 1990. Peptide binding to HLA-DR1: a peptide with most residues substituted to alanine retains MHC binding. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:1797.

61. Rognan, D., M.J. Reddehase, U.H. Koszinowski, and G. Folkers. 1992. Molecular modeling of an antigenic complex between a viral peptide and a class I major histocompatibility glycoprotein. *Proteins Struct. Funct. Genet.* 13:70.

62. DeLisi, C., and J.A. Berzofsky. 1985. T-cell antigenic sites tend to be amphipathic structures. *Proc. Natl. Acad. Sci. USA.* 82:7048.

63. Margalit, H., J.L. Spouge, J.L. Cornette, K.B. Cease, C. DeLisi, and J.A. Berzofksy. 1987. Prediction of immunodominant helper T cell antigenic sites from the primary structure. *J. Immunol.* 138:2213.

64. Takahashi, H., J. Cohen, A. Hosmalin, K.B. Cease, R. Houghten, J.L. Cornette, C. DeLisi, B. Moss, R.N. Germain, and J.A. Berzofsky. 1988. An immunodominant epitope of the human immunodeficiency virus envelope glycoprotein gp160 recognized by class I major histocompatibility complex molecule-restricted murine cytotoxic T lymphocytes. *Proc. Natl. Acad. Sci. USA.* 85:3105.

65. Reyes, V.E., E.J. Fowlie, S. Lu, L. Phillips, L.T. Chin, R.E. Humphreys, and R.A. Lew. 1990. Comparison of three related methods to select T cell-presented sequences of protein antigens. *Mol. Immunol.* 27:1021.

66. Allen, P.M., G.R. Matsueda, R.J. Evans, J.B. Dunbar, G.R. Marshall, and E.M. Unanue. 1987. Identification of the T-cell and Ia contact residues of a T-cell antigenic epitope. *Nature (Lond.).* 327:713.

67. Sette, A., S. Buus, S. Colon, J.A. Smith, C. Miles, and H.M. Grey. 1987. Structural characteristics of an antigen required for its interaction with Ia and recognition by T cells. *Nature (Lond.).* 328:395.

68. Abergel, C., E. Loret, and J.M. Claverie. 1989. Conformational analysis of T immunogenic peptides by circular dichroism spectroscopy. *Eur. J. Immunol.* 19:1969.

69. Fersht, A.R., J.P. Shi, J. Knill-Jones, D.M. Lowe, A.J. Wilkinson, D.M. Blow, P. Brick, P. Carter, M.M. Waye, and G. Winter. 1985. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature (Lond.).* 314:235.

70. Bordo, D., and P. Argos. 1990. Evolution of protein cores, constraints in point mutations as observed in globin tertiary structures. *J. Mol. Biol.* 211:975.

71. Udaka, K., T.J. Tsomides, and H.N. Eisen. 1992. A naturally occurring peptide recognized by alloreactive CD8+ cytotoxic T lymphocytes in association with a class I MHC protein. *Cell.* 69:989.

72. Reddehase, M.J., and U.H. Koszinowski. 1991. Redistribution of critical major histocompatibility complex and T cell receptor-binding functions of residues in an antigenic sequence after biterminal substitution. *Eur. J. Immunol.* 21:1697.

73. Bordo, D., and P. Argos. 1991. Suggestions for "safe" residue substitutions in site-directed mutagenesis. *J. Mol. Biol.* 217:721.

74. Thomas, M.L., P.J. Reynolds, A. Chain, Y. Ben-Neriah, and I.S. Trowbridge. 1987. B-cell variant of mouse T200 (Ly-5): evidence for alternative mRNA splicing. *Proc. Natl. Acad. Sci. USA.* 84:5360.

75. Marquardt, H., G.J. Todaro, and M. Shoyab. 1986. Complete amino acid sequences of bovine and human endozepines. Homology with rat diazepam binding inhibitor. *J. Biol. Chem.* 261:9727.

76. Uchiyama, H., and B. Weisblum. 1985. N-Methyl transferase of *Streptomyces erythraeus* that confers resistance to the macrolide-lincosamide-streptogramin B antibiotics: amino acid sequence and its homology to cognate R-factor enzymes from pathogenic bacilli and cocci. *Gene.* 38:103.

77. Ann, D.K., and D.M. Carlson. 1985. The structure and organization of a proline-rich protein gene of a mouse multigene family. *J. Biol. Chem.* 260:15863.

78. Guillemot, F., A. Billault, O. Pourquié, G. Béhar, A.M. Chaussé, R. Zoorob, G. Kreibich, and C. Auffray. 1988. A molecular map of the chicken major histocompatibility complex: the class II β genes are closely linked to the class I genes and the nucleolar organizer. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:2775.

79. Brock, C.J., and J.E. Walker. 1980. Superoxide dismutase from *Bacillus stearothermophilus.* Complete amino acid sequence of a manganese enzyme. *Biochemistry.* 19:2873.

80. Nudel, U., K. Robzyk, and D. Yaffe. 1988. Expression of the putative Duchenne muscular dystrophy gene in differentiated myogenic cell cultures and in the brain. *Nature (Lond.).* 331:635.

81. Adra, C.N., P.H. Boer, and M.W. McBurney. 1987. Cloning and expression of the mouse pgk-1 gene and the nucleotide sequence of its promoter. *Gene.* 60:65.

82. Connolly, M.L. 1983. Analytical molecular surface calculation. *J. Appl. Crystallog.* 16:548.

83. DeLean, A., P.J. Munson, and D. Rodbard. 1978. Simultaneous analysis of families of sigmoidal curves: application to bioassay, radioligand assay, and physiological dose–response curves. *Am. J. Physiol.* 235:E97.

84. Padlan, E.A. 1990. On the nature of antibody combining sites: unusual structural features that may confer on these sites an enhanced capacity for binding ligand. *Proteins Struct. Funct. Genet.* 7:112.