Crystal Structures of Two Viral Peptides in Complex with Murine MHC Class I H-2K^b

Daved H. Fremont,* Masazumi Matsumura, Enrico A. Stura, Per A. Peterson, Ian A. Wilson†

The x-ray structures of a murine MHC class I molecule (H-2K^b) were determined in complex with two different viral peptides, derived from the vesicular stomatitis virus nucleoprotein(52–59), VSV-8, and the Sendai virus nucleoprotein(324–332), SEV-9. The H-2K^b complexes were refined at 2.3 Å for VSV-8 and 2.5 Å for SEV-9. The structure of H-2K^b exhibits a high degree of similarity with human HLA class I, although the individual domains can have slightly altered dispositions. Both peptides bind in extended conformations with most of their surfaces buried in the H-2K^b binding groove. The nonamer peptide maintains the same amino- and carboxyl-terminal interactions as the octamer primarily by the insertion of a bulge in the center of an otherwise β conformation. Most of the specific interactions are between side-chain atoms of H-2K^b and main-chain atoms of peptide. This binding scheme accounts in large part for the enormous diversity of peptide sequences that bind with high affinity to class I molecules. Small but significant conformational changes in H-2K^b are associated with peptide binding, and these synergistic movements may be an integral part of the T cell receptor recognition process.

Molecular recognition processes are fundamental to both humoral and cellular immunity. Despite similarities in genetics and structure (*1*, *2*), antibodies and T cell receptors (TCR's) recognize foreign antigens differently. Antibodies bind free antigens while TCR's recognize the universe of antigenic peptides presented by major histocompatibility complex (MHC) molecules of the class I and class II types (*3*–*5*). However, many peptides bound to MHC molecules may not be recognized by TCR's, in part because cells expressing TCR that are reactive with peptides derived from the individual's own proteins are eliminated during T cell development in the thymus. The population of MHC molecules expressed in an individual is extremely limited compared to the diverse repertoire of the TCR. Nevertheless, MHC molecules bind and present a vast array of foreign and self-derived peptides while TCR molecules are restricted to recognizing an extremely limited subset. This situation appears paradoxical in that MHC class I molecules bind peptides with an affinity at least 1000 times greater than that of TCR binding to antigen (*6*).

Class I MHC-peptide complexes are recognized by cytotoxic T lymphocytes, which use the accessory class I binding CD8 molecule to focus the TCR onto the antigenic structure (*1*–*5*). Crystallographic evidence (*7*–*8*) reveals that class I molecules contain two membrane proximal, immunoglobulin-like domains. Distal to the membrane are eight antiparallel β strands and two long α helices that form the bottom and sides of the peptide binding groove. Although the shape of the groove has been resolved for three human class I molecules (*7*–*11*), the precise nature of the interactions between MHC residues in the groove and bound peptides has resisted analysis. The major reason for this difficulty has been the unavailability of class I molecules with a homogeneous peptide content (*12*–*15*). Nevertheless, crystallographic studies (*11*) of a human class I molecule in complex with a mixture of peptides provided some key features of peptide binding.

To address the issues of how a single class I molecule can bind a vast number of chemically distinct peptides and how TCR's recognize class I peptide complexes, we have devised procedures to generate homogeneous class I peptide complexes (*16*, *17*) and have succeeded in crystallizing them (*18*). We have now determined the structure of a murine class I molecule, H-2K^b, and present the structural analysis of an MHC class I molecule bound to two specific peptide antigens. The peptides, which are of different lengths, are derived from the vesicular stomatitis virus nucleoprotein(52–59), RGYVYQGL (VSV-8), (*19*, *20*) and Sendai virus nucleoprotein(324–332), FAPGNYPAL (SEV-9) (*21*). These peptides, which bind with nanomolar affinities (*17*), contain residues consistent with the proposed H-2K^b–restricted peptide motif (*13*), with the anchor tyrosine at position 5 or 6 and a leucine at position 8 or 9. We also describe in detail the molecular basis of recognition of antigenic peptides by MHC allele-specific class I molecules (*22*).

Structure determination: The structure of murine MHC class I H-2K^b was determined by molecular replacement (MR). Initially, H-2K^b–VSV-8 crystals were grown in a triclinic space group, *P*1, with two molecules per asymmetric unit and diffracted to only medium resolution (~3 Å). The MR solution was obtained with the use of coordinates of HLA-A2 (*9*) as the probe model. Although preliminary structure refinement proceeded smoothly, electron density maps, both averaged and unaveraged, were not of exceptional quality. Before refinement was completed, better diffracting crystals of both VSV-8 and SEV-9 complexes were obtained in the orthorhombic space group $P2_12_12$, with a single molecule per asymmetric unit (*18*). Only the orthorhombic structures are presented here (Table 1).

An MR solution for the H-2K^b–SEV-9 complex was determined with the partially refined triclinic H-2K^b–VSV-8 coordinates. After one cycle of refinement by simulated annealing, the H-2K^b–SEV-9 electron density map showed interpretable density for both the heavy chain (residue numbering 1–274) and β_2-microglobulin (β_2m, B1-99). The electron density was at first difficult to interpret for a few of the loops connecting the β strands and at the amino and COOH-termini of both chains. However, after seven more cycles of refinement and manual rebuilding, the electron density maps were of excellent quality. The trace of the main chain was unambiguous with the exception of heavy chain residues 219–227, corresponding to a surface-exposed loop with extremely weak density. Otherwise, good density was observed for the heavy chain up to residue 272, with weak but interpretable density for residues 273–274. No interpretable density was seen for the carbohydrate expected to be linked to Asn^{86}, which is solvent exposed. Clear density was seen for β_2m with the exception of the NH_2-terminal Ile^{B1}, which appears slightly disordered. Continuous electron density in the binding groove corresponding to the SEV-9 peptide (P1–9) was evident from the first maps, although its quality was not sufficient for thorough interpretation. After a complete round of building and a second cycle of simulated annealing refinement with H-2K^b coordinates only, the entire SEV-9 could be traced in an F_o

D. H. Fremont, E. A. Stura, and I. A. Wilson are in the Department of Molecular Biology and M. Matsumura and P. A. Peterson are in the Department of Immunology at the Scripps Research Institute, La Jolla, CA 92037.

*D.H.F. is a graduate student in the Department of Chemistry, University of California–San Diego, La Jolla, CA 92093.
†To whom correspondence should be addressed.

$- F_c$ difference map. However, coordinates for SEV-9 were not actually included in refinement until the R value had dropped below 20 percent for 8.0 to 2.6 Å data. The final SEV-9 model has excellent geometry and contains 91 water molecules, giving an R value of 0.154 to 2.5 Å resolution (Table 1).

Refinement of the VSV-8 complex at 2.3 Å then began with the partially refined H-2K^b–SEV-9 coordinates, and clear, continuous density for the peptide (P1–P8) was evident in the initial difference map. The VSV-8 coordinates were not included in refinement until the fifth cycle of rebuilding of the H-2K^b model. Regions of H-2K^b that were difficult to interpret in the SEV-9 complex were also generally unclear in the VSV-8 complex, and information from one map was consulted to aid in the interpretation of the other. The final VSV-8 model contains 114 water molecules and has an R value of 0.169 to 2.3 Å resolution (Table 1).

Structural comparison with human MHC class I. The overall structure of H-2K^b is similar to HLA class I molecules (*7–11*) (Fig. 1). The two H-2K^b polypeptide chains fold into three distinct domains: the two membrane distal domains, encoded by exons α_1(residues 1–90) and α_2(91–182), are related by a pseudo-twofold rotation axis and form a superdomain to assemble the β sheet and α-helix platform for peptide binding. The α_3(183–274) and β_2m(B1–99) domains interact in nonstandard pairings compared to immunoglobulins and are located beneath the $\alpha_1\alpha_2$ domain (Fig. 1). Each related domain of H-2K^b and HLA-A2 is similar in both sequence and structure. The individual $\alpha_1\alpha_2$, α_3, and β_2m domains share 70 to 71 percent sequence identity. Superimposition (*23*) of the $\alpha_1\alpha_2$ domains of H-2K^b–VSV-8 and HLA-A2 gives a root-mean-square (rms) deviation for Cα positions of 0.66 Å (compared to 0.32 Å for the two H-2K^b complexes). The α_3 and β_2m domains have larger rms deviations of 1.13 and 0.81 Å, respectively (compared to 0.30 and 0.23 Å for the two H-2K^b complexes).

Comparison of the secondary structure of H-2K^b with HLA-A2 yields some interesting observations. In the $\alpha_1\alpha_2$ domain, all eight antiparallel strands that make up the peptide binding platform are identical [for secondary structure nomenclature, see (*9*)]. However, the solvent exposed loops connecting the strands show slight variation, most obviously in the connection from S1 to S2 (13–20) and from S3 to S4 (38–40) in α_1. In HLA-A2, residues 50 to 53 form a single turn of α helix between strand S4 and the long helix (57–85), which forms one wall of the peptide binding groove. In H-2K^b, residues 50 to 54 form a 3_{10} helix while the long helical segment (57–84) is conserved. The α-helical region of α_2, which forms the other side of the binding groove, also shows secondary structure similarity in HLA and H-2K^b. In H-2K^b, the extended helix is broken three times by glycine into four shorter helical segments, H1 (138–150), H2a (152–161), H2b (163–174), and H3 (176–179). The HLA sequences lack Gly^{151}, but nevertheless have the same segments. The H1 helices of the VSV-8 and SEV-9 complexes

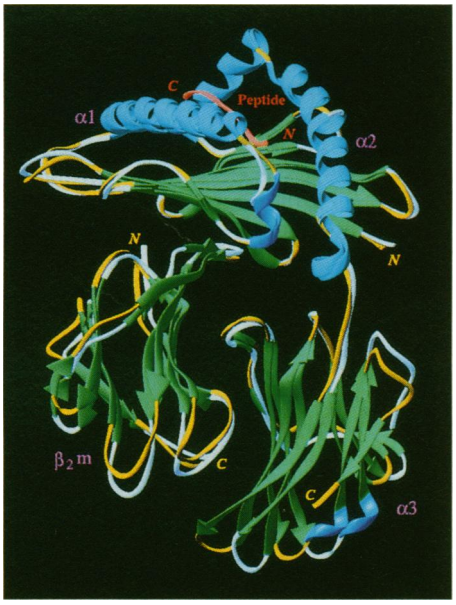

Fig. 1. Comparison of the murine and human MHC class I molecules. A backbone ribbon representation (*49*) of the murine H-2K^b–VSV-8 complex is shown superimposed onto HLA-A2. The Cα coordinates of only the $\alpha_1\alpha_2$ domains were used for the alignment in order to accentuate the differences in dispositions of the α_3 and β_2m domains. For both structures, β strands are shown in green, α-helices in cyan, and 3_{10} helices in blue. Connecting loops of H-2K^b are shown in yellow, HLA-A2 in white. The VSV-8 peptide is red. The main-chain difference in human and murine β_2m strand S4 can be seen directly below the α_1 domain.

Table 1. Diffraction data and crystallographic statistics for the H-2K^b orthorhombic ($P2_12_12$) VSV-8 and SEV-9 complexes. Soluble, peptide-free H-2K^b molecules were expressed in *Drosophila* cells and prepared as single-peptide complexes (*17*). Data collection quality crystals were grown by sitting drop vapor diffusion at 22.5°C by first equilibrating equal volume drops consisting of H-2K^b-peptide complex (4 to 6 mg/ml) solution and 1.84 M NaH_2PO_4, K_2HPO_4, solution with 2 percent MPD, pH 6.25, for 2 to 3 days with subsequent macroseeding of the drop with previously grown crystals. Further details of the crystallization of these and other crystal forms, including a third complex with OVA-8 (SIINFEKL) (*42*), have been described (*18*). Data were collected on a Siemens proportional counter multiwire area detector with an Elliott GX-18 rotating anode x-ray generator (40 kV and 55 mA) with a 100-μm focal cup and Franks focusing mirrors. Data were indexed and reduced by means of XENGEN (*43*). The reported data are from one crystal of VSV-8 and two crystals of SEV-9. The molecular replacement structure of murine MHC class I H-2K^b was solved initially in space group $P1$ with the use of coordinates of HLA-A2 [PDB entry 3HLA (*9*)] as the probe model (*31*). The partially refined coordinates from the triclinic crystals served as the model for the molecular replacement structure solution of the orthrhombic crystals. Self and cross-rotation studies were done with MERLOT (*44*). Translation functions were undertaken with HARADA, a fast Fourier-based full cell search program (*45*). Rigid body, least squares, and simulated annealing refinement were all performed with XPLOR 2.1 (*46*) on a Cray-YMP supercomputer. Refinement was initiated by rigid body refinement treating the $\alpha_1\alpha_2$, α_3, and β_2m domains as separate pieces. Simulated annealing refinements were performed with temperatures as high as 5000 K. Electron density maps were visualized with FRODO (*47*) on an Iris 310 work station. Model building was done primarily using $F_o - F_c$ omit maps calculated by removing 30 residues of the model for 50 cycles of positional refinement and for the subsequent structure factor calculation. Coordinate errors are estimated (*48*) to be between 0.20 Å and 0.25 Å for both peptide complexes.

	H-2K^b–VSV-8 complex	H-2K^b–SEV-9 complex
Cell constants (Å)	a = 138.1, b = 88.9, c = 45.7	a = 139.0, b = 90.4, c = 46.3
Total observations (maximum resolution)	81,010 (2.3 Å)	78,975 (2.5 Å)
Unique reflections with I ≥ 0 (%)	23,289 (90%)	16,934 (81%)
Unique reflections with I ≥ 2σ (%)	20,624 (80%)	15,513 (74%)
R_{sym} (I)	0.079	0.108
Protein + peptide atoms	3,053 + 68	3,053 + 68
Water molecules	114	91
Refinement resolution	6.0 – 2.3 Å	6.0 – 2.5 Å
rms deviation from ideal bond lengths	0.012 Å	0.012 Å
rms deviation from ideal bond angles	3.06°	3.09°
R value for I ≥ 2σ	0.169	0.154

exhibit small but significant topological differences as a result of their interactions with bound peptide.

The α_3 and β_2m domains are similar in folding topology to an immunoglobulin constant heavy domain and are each composed of two antiparallel β sheets. Some significant differences mark the H-2K^b and HLA-A2 α_3 domains. Whereas strand S7 (270–272) is not observed in the human class I structures, it is clearly seen in H-2K^b. One of the largest local structural difference between HLA-A2 and H-2K^b is in the exposed turn (195–197) connecting S1 to S2 just below the interface with β_2m, where main chain deviations of more than 3 Å are observed. In the H-2K^b structure, strand S1 is extended by one residue (Arg^{194}) and forms a tighter turn than in HLA-A2 because of Pro^{195}, which is rare in both human and murine alleles. An important structural difference between HLA-A2 and H-2K^b occurs in the α_3 surface-exposed residues (223–228) that have been implicated in the CD8 binding of class I molecules (*24*). Although the density is weak, the conformation of this region appears different from that observed for HLA-A2. Main chain differences of up to 2 Å are apparent for the turn proceeding α_3 strand S4 (225–228). Nevertheless, the surface exposure of several negatively charged side chains is maintained in this region between the human and murine structures.

The structures of human and murine β_2m are nearly identical except for two differences. The S4 β strand, which in human β_2m is broken by a β bulge (B52–54), is a continuous β strand (B50–56) in H-2K^b. Only two residues differ in this region between human (His^{B51} and Leu^{B54}) and murine (Met^{B51} and Met^{B54}) β_2m. The other significant main chain difference occurs at the CCOH-terminus, where an altered turn is adopted. Both of these changes lie within regions of murine β_2m that form the interface with α_3 and consequently profoundly affect the interdomain contacts.

Domain structure. The overall assembly of the three domains of murine H-2K^b is considerably different from that in HLA-A2. The geometrical arrangements and domain dispositions are altered such that some interdomain contacts are conserved, but many others are not. In HLA-A2, β_2m and α_3 are related by a 144° rotation and a 14.6 Å translation, whereas in H-2K^b the domains are related by a 159° rotation and a 13.1 Å translation (*25*). In addition, the axis relating the two immunoglobulin-like domains is moved 4° and 9.2 Å relative to the axis relating α_1 and α_2. Neither the two murine nor the human molecules (HLA-A2 and HLA-Aw68) themselves differ in their domain dispositions, although similarity in crystal packing may be a factor.

The contacts of α_3 with $\alpha_1\alpha_2$ are well conserved between H-2K^b and HLA-A2. In both mice and humans, Tyr^{209} plays a central role in this interface by forming a hydrogen bond with the hydroxyl group of Thr^{31}. In addition, a conserved water molecule mediates a hydrogen bonding network connecting the hydroxyl oxygen of Tyr^{209}, the carbonyl oxygen of Leu^{179}, and the carboxyl of Asp^{29}. The only interaction not observed in H-2K^b is a hydrogen bond between the Tyr^{209} hydroxyl and the amide nitrogen of Asp^{30}, although the side chain of Asp^{30} still maintains its hydrogen bond with the backbone amide of Ala^{211}.

It has been well established that β_2m stabilizes the structure of the class I heavy chain (*16*, *26–28*). The size of the β_2m-$\alpha_1\alpha_2$ interface is similar for human and murine class I molecules, with an average of 520 and 550 Å^2 of buried surface (*29*) for β_2m and $\alpha_1\alpha_2$, respectively. However, the specific interactions of β_2m with $\alpha_1\alpha_2$ differ more than the sequence similarity at the interface would imply. Of the 28 residues that contribute to the interface, only seven are different between the two species (23, 32, 98, 121, B33, B34, and B54). Of the nine hydrogen bonds at the interface of $\alpha_1\alpha_2$ and β_2m in H-2K^b (Table 2), only

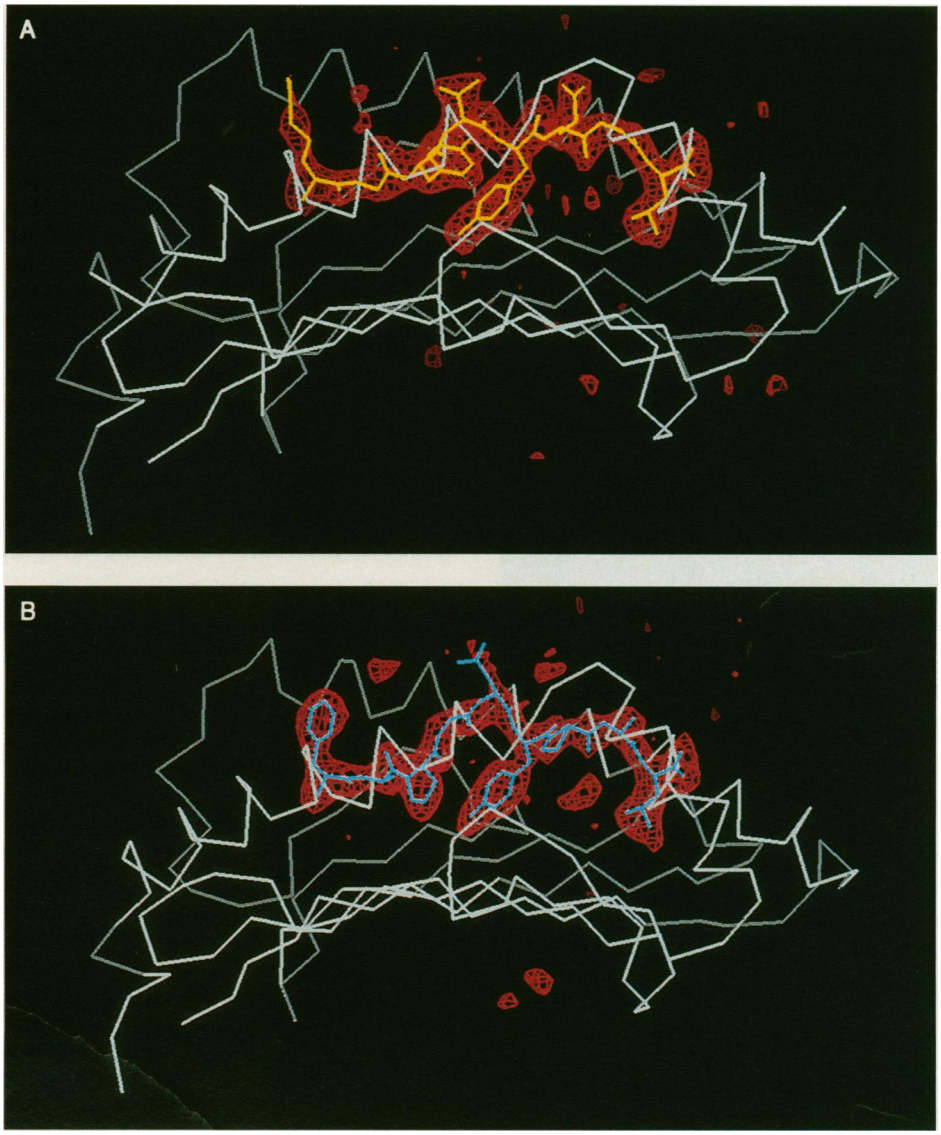

Fig. 2. Electron density for the bound VSV-8 (**A**) and SEV-9 (**B**) peptides. The $F_o - F_c$ omit electron density maps were contoured at 3.0 σ and were calculated from phases in which 40 cycles of positional refinement were undertaken in the absence of the peptide coordinates to remove model bias using XPLOR (*46*) and displayed (in red) using FRODO (*47*). The Cα trace of the H-2K^b $\alpha_1\alpha_2$ domain is shown in white. In this and subsequent figures, the VSV-8 peptide is displayed in yellow and the SEV-9 is displayed in cyan. The NH_2-terminus of each peptide is on the left and their orientation is parallel with the α_1 flanking helix, which is seen behind the peptide.

three are conserved in HLA-A2. On the other hand, the extensive hydrophobic interactions of Phe^{B56} and Trp^{B60} at the interface are conserved. Most of the interdomain contact differences are the direct result of the altered main-chain structure of β-strand S4 of murine β_2m. For instance, the carboxylate of Asp^{B53}, which in HLA-A2 forms a salt bridge with Arg^{35}, has a completely different location in H-2K^b (shifted 8 Å toward the α_3 interface). Four hydrogen bonds are formed with the main-chain of β-strand S4, and indeed, the possibility exists that the observed conformation is stabilized by its interactions with the H-2K^b heavy chain. The only significant conformational difference seen in $\alpha_1\alpha_2$ is the side chain of Arg^{35}, which adopts different rotamers in the human and murine interfaces.

The differing interactions at the interface of β_2m and α_3 between HLA-A2 and H-2K^b are primarily due to the differences in domain dispositions. The interface is slightly larger for the murine heterodimer with 495 Å^2 and 520 Å^2 of buried surface for the α_3 and β_2m domains compared to 15 and 9 percent less in HLA-A2. Of 13 hydrogen bonds between β_2m and α_3, only four are conserved (Table 2). Although the conformation of α_3 residues at the interface is fairly similar between HLA-A2 and H-2K^b, significant differences are observed in β_2m. For instance, Tyr^{B63}, which in HLA-A2 interacts with Tyr^{27}, is rotated into the α_3 interface, and Arg^{B12} adopts a different rotamer as a result of the rigid body changes. The greatest structural difference occurs at the COOH-terminus of β_2m, which, although tightly packed in both interfaces, is completely reoriented in H-2K^b resulting in substantially altered interactions.

Human β_2m not only binds to H-2K^b, but also increases its thermal stability (*16*, *26*). Docking experiments with coordinates of human β_2m and H-2K^b indicate that only one significant steric overlap (Asp^{B53}–Arg^{35}) would occur if β_2m bound with either the human or murine domain dispositions. Indeed, this interaction could be favorable with an alteration of side-chain rotamers. Whether the differences in the β_2m structure are a result of its interaction with heavy chain is not known.

Peptide conformation. Both the VSV-8 and the SEV-9 peptides bind into the deep groove of the $\alpha_1\alpha_2$ domain. The peptides are in extended conformations, stretching 25.1 and 25.4 Å from NH_2- to COOH-terminus for VSV-8 and SEV-9, respectively. The final electron density maps (Fig. 2, A and B) for both peptides show no main chain breaks and only minor ambiguities. For the VSV-8 peptide, there are two ambiguities worth noting. Electron density for the side chain of Arg^{P1}, which projects out of the binding cleft, is rather weak past Cδ. The side chain of Gln^{P6}, which in our model is built facing the solvent, possibly has a less populated secondary conformation in which the side chain points down toward the β-sheet platform. Similarly, the electron density for SEV-9 is of excellent quality except for the highly exposed Asn^{P5}.

The overall conformations of the two peptides are similar (Fig. 3). Residues in both peptides have backbone torsion angles corresponding to the β region of phi-psi space except for SEV-9 Asn^{P5}, which adopts an alpha left conformation ($\phi = 42°$, $\psi = 44°$). Both prolines of the SEV-9 peptide are in the trans conformation. The additional residue in SEV-9 is primarily accommodated as an insertion in the center of the sequence, with residues Gly^{P4} and Asn^{P5} protruding from the binding site as a β bulge in roughly the same position as Val^{P4} of VSV-8 (Fig. 3). Indeed, the Cα position of VSV-8 Val^{P4} is only 1.4 Å and 2.9 Å from the Cα of SEV-9 Gly^{P4} and Asn^{P5}, respectively (these and all subsequent distances are

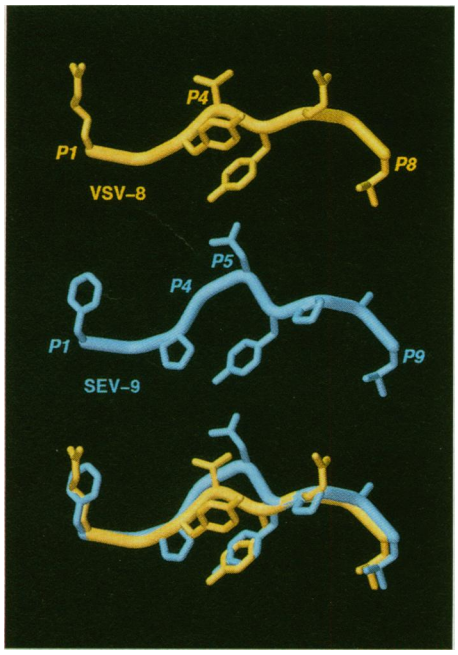

Fig. 3. Comparison of the VSV-8 and SEV-9 peptide conformations. The Cα traces of the peptide backbones are shown as thick tubes (*50*). VSV-8 (RGYVYQGL) is shown in yellow, SEV-9 (FAPGNYPAL) in cyan. The peptides were oriented by superimposing only the Cα atoms of the $\alpha_1\alpha_2$ domains of the two H-2K^b complexes. The overlap of the two peptides (bottom) shows that the NH_2- and COOH-terminal residues have almost identical conformations while the extra residue in the nonamer (P5) is accommodated in a β-bulge.

Table 2. Hydrogen bonds formed between murine β_2m and the heavy chain of H-2K^b. Hydrogen bonds <3.4 Å are presented (*23*). Conserved interactions with HLA-A2 (*9*) are highlighted in bold. *HisB^{34} appears to adopt two conformations, only one of which interacts with Ser^{92}.

β_2m - $\alpha_1\alpha_2$ interface		β_2m - α_3 interface	
β_2m	$\alpha_1\alpha_2$	β_2m	α_3
I^{B1}(N)	**D^{119}(O)**	**Q^{B8}(Oε1)**	**R^{234}(Nη2)**
H^{B31}(Nε2)	**Q^{96}(Oε1)**	**Y^{B10}(OH)**	**P^{235}(O)**
H^{B34}(Nε2)*	**S^{92}(Oγ)**	**S^{B11}(O)**	**Q^{242}(Nε2)**
M^{B51}(O)	**R^{35}(Nη1)**	**R^{B12}(Nε)**	**D^{238}(Oδ2)**
D^{B53}(N)	**E^{32}(Oε2)**	**R^{B12}(Nε)**	**T^{240}(Oγ1)**
M^{B54}(N)	**Y^{27}(OH)**	**N^{B24}(Nδ2)**	**A^{236}(O)**
M^{B54}(O)	**Y^{27}(OH)**	**Q^{B29}(Nε2)**	**E^{232}(Oε1)**
W^{B60}(Nε1)	**D^{122}(Oδ2)**	**Y^{B63}(OH)**	**E^{232}(Oε1, 2)**
W^{B60}(O)	**Q^{96}(Nε2)**	**M^{B99}(O)**	**T^{190}(Oγ1)**
		M^{B99}(O)	**H^{192}(Nε2)**
		M^{B99}(O)	**R^{202}(Nε)**
		M^{B99}(OT)	**W^{204}(Nε1)**

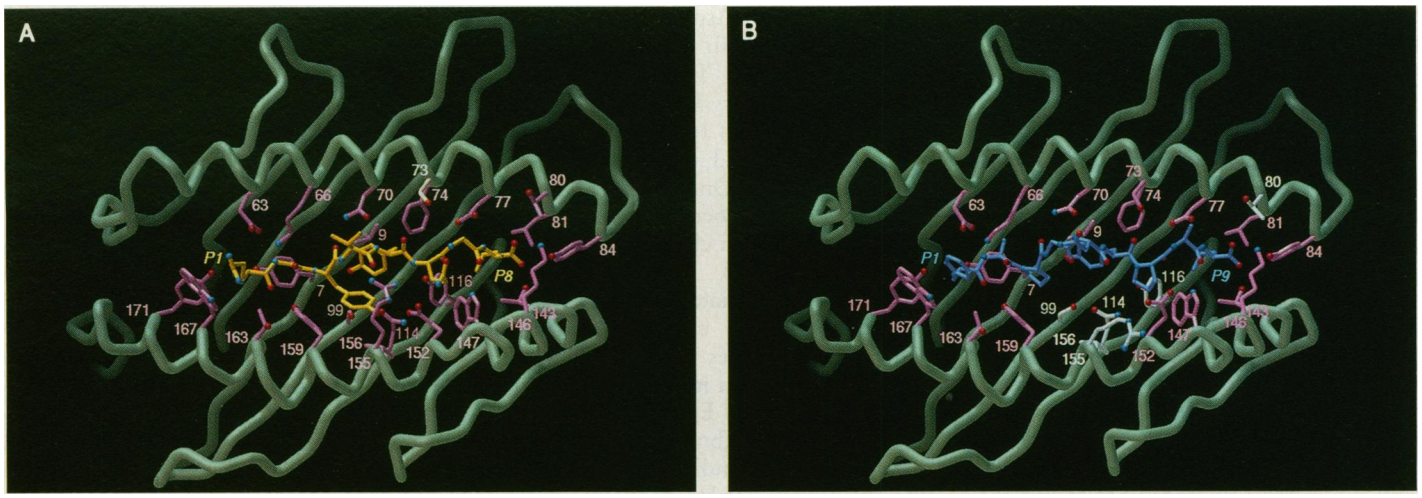

Fig. 4. Interactions of the two viral peptides with H-2K^b. The H-2K^b residues in van der Waals contact with both the VSV-8 (**A**) and SEV-9 (**B**) peptides are shown. The Cα backbone of the $\alpha_1\alpha_2$ domain is represented as a gray tube (*50*). Side chains in contact (<4.0 Å) with a given peptide are displayed in pink, and side chains which contact one peptide but not the other are shown in gray (Ser^{73} for VSV-8; Thr^{80}, Ser^{99}, Gln^{114}, Tyr^{116}, Arg^{155}, and Leu^{156} for SEV-9). The VSV-8 peptide is yellow, the SEV-9 peptide is cyan. Side chain and peptide oxygens and nitrogens are represented as red and blue balls, respectively.

based only on superimposed Cα coordinates of $\alpha_1\alpha_2$). The additional residue also results in a slight elongation of the COOH-terminus of SEV-9 compared to VSV-8 (~0.6 Å). The backbone conformations of the first three and last four residues in each peptide are similar, with an average rms deviation in dihedral angles of 19.9°. The corresponding overlap of the Cα atoms shows an rms deviation of 0.71 Å, the largest difference (0.97 Å) at the H-2K^b-restricted anchor residue VSV-8 Tyr^{P5} and SEV-9 Tyr^{P6}. The side-chain positions and rotamers are likewise conserved. For the anchor tyrosine residue, the side-chain torsion angles deviate by no more than 11° (χ_1 = 61° and χ_2 = −85° in VSV-8); for the other conserved residue VSV-8 Leu^{P8} and SEV-9 Leu^{P9}, they deviate by no more than 4° (χ_1 = −57° and χ_2 = −170° in VSV-8). The hydroxyl oxygens of the anchor tyrosines are 1.2 Å apart from one another, but their water-mediated interactions with H-2K^b are mostly conserved (see below). The NH_2-terminal VSV-8 Arg^{P1} and SEV-9 Phe^{P1} share a similar χ_1, and indeed, the superimposed peptide coordinates show the aliphatic chain of Arg^{P1} projecting through the phenyl ring of Phe^{P1} (Fig. 3).

Table 3. Van der Waals interactions between peptide and H-2K^b. Contacts <4.0 Å between peptide and H-2K^b are listed. Also presented are the number of bound waters in contact with peptide. The number of contacts are shown in parentheses.

VSV-8 interactions		SEV-9 interactions	
R^{P1}	Y^7 (5), E^{63} (3), K^{66} (1), Y^{159} (3), T^{163} (1), W^{167} (12), Y^{171} (4), 2 H$_2$O(4)	F^{P1}	Y^7 (5), E^{63} (2), K^{66} (3), Y^{159} (3), T^{163} (2), W^{167} (16), Y^{171} (5), 2 H$_2$O (4)
G^{P2}	Y^7 (4), E^{63} (1), K^{66} (2), Y^{159} (1), 1 H$_2$O (1)	A^{P2}	Y^7 (4), E^{63} (3), K^{66} (3), 2 H$_2$O (2)
Y^{P3}	N^{70} (1), E^{152} (2), R^{155} (9), L^{156} (2), Y^{159} (8), 2 H$_2$O (4)	P^{P3}	N^{70} (1), Y^{159} (3), 1 H$_2$O (2)
-		G^{P4}	1 H$_2$O (1)
V^{P4}	K^{66} (1), N^{70} (1), R^{155} (3)	N^{P5}	2 H$_2$O (4)
Y^{P5}	V^9 (3), N^{70} (5), F^{74} (4), S^{99} (1), Q^{114} (2), Y^{116} (1), 6 H$_2$O (11)	Y^{P6}	V^9 (1), N^{70} (5), S^{73} (1), F^{74} (1), 5 H$_2$O (8)
Q^{P6}	W^{147} (3), E^{152} (10), R^{155} (2), 3 H$_2$O (5)	P^{P7}	W^{147} (4), E^{152} (3)
G^{P7}	D^{77} (1), W^{147} (1), 3 H$_2$O (3)	A^{P8}	D^{77} (2), W^{147} (2), 1 H$_2$O (1)
L^{P8}	D^{77} (5), T^{80} (1), L^{81} (1), Y^{84} (5), Y^{116} (1), T^{143} (6), K^{146} (2),W^{147} (1), 2 H$_2$O (2)	L^{P9}	D^{77} (3), L^{81} (3), Y^{84} (4), T^{143} (4), K^{146} (3), W^{147} (2), 1 H$_2$O (2)
Totals:	119 contacts with H-2K^b 30 contacts with water		88 contacts with H-2K^b 24 contacts with water

Peptide interactions. Probably the most striking feature of these structures is the great depth at which the peptides are buried in the H-2K^b antigen binding groove. For VSV-8, 681 Å^2 (83 percent) of peptide surface area becomes inaccessible to solvent upon binding (*30*), while 846 Å^2 (composed of 37 residues) of the H-2K^b surface is buried. For SEV-9, 604 Å^2 (75 percent) of peptide and 820 Å^2 (36 residues) of H-2K^b are buried. The difference between the two buried peptide surfaces can be primarily attributed to the bulge in the center of SEV-9. Only a few of the individual residues are solvent exposed in the complexes (*22*, figure 5A). A total of 24 (23 in VSV-8, 18 in SEV-9) residues from H-2K^b are in van der Waals contact with peptide, 11 from α_1 (9 from helix, 2 from sheet) and 13 from α_2 (10 from helix, 3 from sheet). In VSV-8 there

are 149 van der Waals contacts, 30 of which are to 14 bound water molecules (Table 3 and Fig. 4A). SEV-9 makes a total of 112 van der Waals contacts, 24 of which are to 9 bound water molecules (Table 3 and Figure 4B). The smaller number of water molecules seen for SEV-9 may in part be due to the current resolution of the structure refinement (2.5 Å compared to 2.3 Å for VSV-8).

Although the peptide main chain and side chain atoms are equally buried in both complexes, 96 percent of the buried H-2K^b atoms are from side chains. Indeed, the most of the specific (electrostatic) interactions are from H-2K^b side chains to the peptide main chain (Table 4). For VSV-8, 14 direct hydrogen bonds and one salt bridge (between the peptide carboxylate and Lys^{146}) are formed between H-2K^b and the peptide main chain. Hydrogen bonds are formed with every carbonyl oxygen and all but two amide nitrogens (VSV-8 Val^{P4} and Gly^{P7}), although hydrogen bonding to four of these atoms is mediated through water molecules. The specific interactions of H-2K^b with SEV-9 are similar to VSV-8, especially at the NH_2 and COOH termini (*22*, figure 4). Deviations are mostly related to the additional length of the SEV-9 peptide as neither Gly^{P4} nor Asn^{P5}, which bulge out of the groove, are in van der Waals contact with H-2K^b. The SEV-9 peptide makes 11 hydrogen bonds and one salt bridge directly with H-2K^b, with all protons either donated or accepted by the peptide backbone. One difference between the two peptides occurs at the anchor tyrosines. The main chain hydrogen bond between VSV-8 Tyr^{P5}(N) and Asn^{70}(Oδ1) is mediated by a water in the SEV-9, while the carbonyl oxygen of SEV-9 Tyr^{P6} makes a direct hydrogen bond with Ser^{73}, an interaction mediated by water in VSV-8. Every hydrogen bond to the peptide backbone of both peptides is formed with (and water bridged to) side-chain atoms of H-2K^b.

The peptide side chains make very few specific interactions. For instance, no side chain from SEV-9 participates in direct hydrogen bonding with H-2K^b. However, the SEV-9 Tyr^{P6} hydroxyl hydrogen bonds with two waters, one of which interacts with Glu^{24}, the other with Ser^{99} and the carbonyl of Tyr^7, the latter being the only interaction with the main-chain of H-2K^b for either peptide. Two VSV-8 side chains hydrogen bond directly with H-2K^b, Tyr^{P3} with Glu^{152} and Arg^{155}, and Gln^{P6} also with Glu^{152}. The only charged residue in either peptide, Arg^{P1}, appears to hydrogen bond with a bound water which in turn interacts with Glu^{63}. The hydroxyl of VSV-8 anchor residue Tyr^{P5} also does not hydrogen bond directly with H-2K^b, but rather is part of a hydrogen bond network similar to that described above for SEV-9, with the substitution of the Tyr^{159} hydroxyl for the Tyr^7 carbonyl. Another similarity between the two peptides is the network of van der Waals interactions of both Arg^{P1} and Phe^{P1} with Trp^{167}. In both complexes, Trp^{167} forms the greatest number of van der Waals interactions with peptide, and can be envisaged as an NH_2-terminal "wall" of the groove, potentially preventing peptides with NH_2-terminal extensions from binding with high affinity.

Peptide associated conformational changes. Small but significant conformational changes in H-2K^b are apparent be-

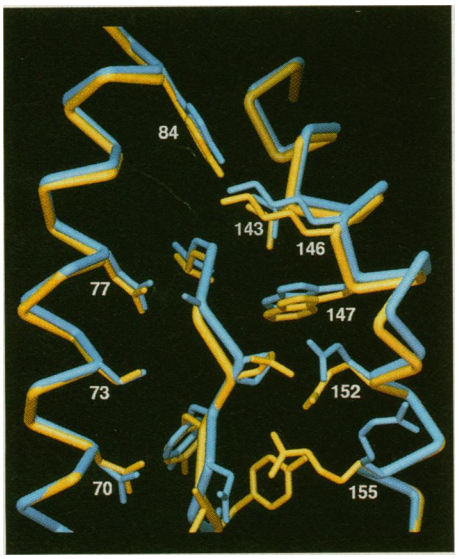

Fig. 5. Peptide-associated conformational changes in H-2K^b. Displayed are Cα tubes (*50*) corresponding to the α helices flanking the COOH-termini of the VSV-8 (yellow) and SEV-9 (cyan) peptides. The H-2K^b side chains that form polar interactions directly with peptide are also shown. The Arg^{155} and Glu^{152} adopt alternate conformations in the two complexes, apparently as a result of the different bound peptide sequences. The main chain of α_2 helix H1 (residues 138–150) is shifted between the two complexes, apparently in order to conserve the interactions of Thr^{143}, Lys^{146}, and Trp^{147} with the peptide backbone.

Table 4. Hydrogen bonds between peptide and H-2K^b. Hydrogen bonds <3.4 Å between peptide and H-2K^b are listed. Those mediated with water molecules are shown in parentheses. Peptide side-chain atoms are shown in bold. Note that Lys^{146} makes a salt bridge.

VSV-8 interactions			SEV-9 interactions		
R^{P1}	-N	Y^7-OH, Y^{171}-OH	F^{P1}	-N	Y^7-OH, Y^{171}-OH
	-O	Y^{159}-OH		-O	Y^{159}-OH
	-Nη2	(E^{63}-Oε2)			
G^{P2}	-N	E^{63}-Oε1, Y^7-OH	A^{P2}	-N	E^{63}-Oε1
	-O	K^{66}-Nζ		-O	K^{66}-Nζ
Y^{P3}	-N	(E^{24}-Oε2, S^{77}-Oγ, Y^{157}-OH)	P^{P3}	-N	
	-O	N^{70}-Nδ2		-O	N^{70}-Nδ2, (E^{24}-Oε2)
	-OH	E^{152}-Oε1, R^{155}-Nε, -Nη1, -Nη2			
-			G^{P4}	-N	
				-O	
V^{P4}	-N		N^{P5}	-N	(N^{70}-Oδ1)
	-O	R^{155}-Nη1, -Nη2		-O	
Y^{P5}	-N	N^{70}-Oδ1	Y^{P6}	-N	(N^{70}-Oδ1)
	-O	(S^{73}-Oγ, N^{70}-Oδ1)		-O	S^{73}-Oγ
	-OH	(E^{24}-Oε2, S^{99}-Oγ, Y159-OH)		-OH	(Y^7-O,E^{24}-Oε2, S^{99}-Oγ)
Q^{P6}	-N	(Y^{116}-OH)	P^{P7}	-N	
	-O	(Y^{116}-OH, D^{77}-Oδ1, S^{73}-Oγ)		-O	
		(R^{155}-Nη1)			
	-Nε2	E^{152}-Oε2			
G^{P7}	-N		A^{P8}	-N	
	-O	W^{147}-Nε1		-O	W^{147}-Nε1
L^{P8}	-N	D^{77}-Oδ2	L^{P9}	-N	D^{77}-Oδ2, (T^{80}-Oγ1)
	-O	Y^{84}-OH, T^{143}-Oγ1		-O	Y^{84}-OH, T^{143}-Oγ1
	-OT	K^{146}-Nζ, (T^{80}-Oγ1, D^{77}-Oδ2)		-OT	K^{146}-Nζ, (T^{80}-Oγ1, D^{77}-Oδ2)
Totals:		**19 direct hydrogen bonds (1 salt bridge) 16 via water**			**11 direct hydrogen bonds (1 salt bridge) 9 via water**

tween the two peptide complexes, although in general the structures show high similarity (the rms deviation for all protein atoms is 0.91 Å). For the current models, only six side-chains in $\alpha_1\alpha_2$ (Asn^{42}, Glu^{46}, Arg^{50}, His^{145}, Glu^{152}, and Arg^{155}) have average rms deviations in excess of 2.0 Å between the two peptide complexes. In both complexes, electron density for the side chains of residues 42, 46, and 50 is relatively weak, as is often the case for polar residues highly exposed to solvent. His^{145} appears to adopt two conformations in both complexes, but so far has been refined with only a single rotamer in each.

The side chains of Glu^{152} and Arg^{155} adopt different conformations as a direct consequence of peptide binding (Figs. 4 and 5). Both of these residues are part of the second helical segment of α_2 (H2a). The larger of these conformational changes is in the side chain location of Arg^{155}, which differs by ~6 Å between the two complexes. In the VSV-8 complex, Arg^{155} forms a salt bridge with Glu^{152} and makes 14 van der Waals contacts with the peptide, including hydrogen bonds to the side chain of Tyr^{P3} and main chain of Val^{P4}. Similar interactions are not formed with SEV-9, because of the peptide bulge and the lack of a polar side chain at P3. Indeed, Arg^{155} makes no contact at all with SEV-9. Instead, the guanidinium group appears to sit near the carboxyl end of the α_2 H1 helix (138–149), forming a salt bridge with Glu^{154} and making hydrogen bonds with the carbonyl oxygens of Ala^{150} and Gly^{151}. The Arg^{155} side chain is far better ordered in the VSV-8 complex than in SEV-9. The other side chain that exhibits a large conformational difference (>3 Å) is Glu^{152}, which has ten van der Waals contacts with VSV-8 and forms hydrogen bonds with the side chains of both Tyr^{P3} and Gln^{P6}. In SEV-9, Glu^{152} has only three contacts, partly because of a 120° χ_1 rotation away from the peptide, and the carboxylate group interacts with a tightly bound water molecule. The electron density for Glu^{152} is completely unambiguous in both the VSV-8 and SEV-9 complexes. In addition to these large conformational differences, several side chains in the peptide binding groove have small torsional differences that alter their positions (<1 Å) in the two complexes, such as for Asn^{70}, Thr^{80}, and Ser^{99}. These three residues participate in hydrogen bonding to the peptide (although only Asn^{70} does so directly) and each seems to move to conserve its interactions with peptide or bound water molecules. Other (smaller) differences between the complexes can be analyzed only with higher resolution refinement of the structures.

Small but systematic main-chain variation is also observed between the two peptide complexes. The backbone shift appears to be a consequence of the slight COOH-terminal extension (0.57 Å) of the SEV-9 peptide relative to VSV-8, as the relative positions of H-$2K^b$ and peptide interacting pairs are highly correlated. In particular, α_2 helix H1 adjusts to accommodate the strongly conserved polar interactions of the peptide carboxylates with Thr^{143} and Lys^{146}, and Trp^{147} with the penultimate carbonyl oxygen (Fig. 5). The observed changes would appear to be close to the error limit of the coordinates. However, as they arise from a concerted shift of the H1 helix from the VSV-8 to SEV-9 conformation, they can be differentiated from random error. The Cα movements average 0.61 Å for residues 142–151. For example, the indole amide nitrogen of Trp^{147} shifts by 0.78 Å, whereas the carbonyl oxygens of SEV-9 Ala^{P8} and VSV-8 Gly^{P7} differ by 0.79 Å. The H-$2K^b$ flexibility is not only confined to the α_2 H1 helix, but the hydroxyl of Tyr^{84} on the long helix of α_1 also shifts by 0.61 Å, apparently to maintain its hydrogen bond with the peptide carboxylate. The positions of Tyr^{84}, Thr^{143}, Trp^{147}, and Tyr^{84} in the SEV-9 complex are all closer to those observed in HLA-A2 (an allele with nonamer peptide preference) than they are to the VSV-8 complex. One of the apparent results of this concerted main chain variation is that the topology of the upper surface of the binding groove is slightly altered, an effect that could have fundamental consequences for TCR recognition of the MHC class I–peptide complex.

Implications for peptide recognition and presentation. Comparison of human and murine MHC class I molecules shows that they adopt the same overall fold but have slightly altered domain dispositions. Like their antibody counterparts in the humoral immune system, the MHC class I molecules also appear to have a flexible elbow or pivot, which in this case connects the $\alpha_1\alpha_2$ domains with α_3 and β_2m (noncovalently). Indeed, we have observed different domain dispositions even within the same peptide-bound H-$2K^b$ in different crystal forms (*31*). In antibodies, the role of elbow flexibility (*32*) has been debated for many years, and the consensus at present is that it does not represent a mechanism for

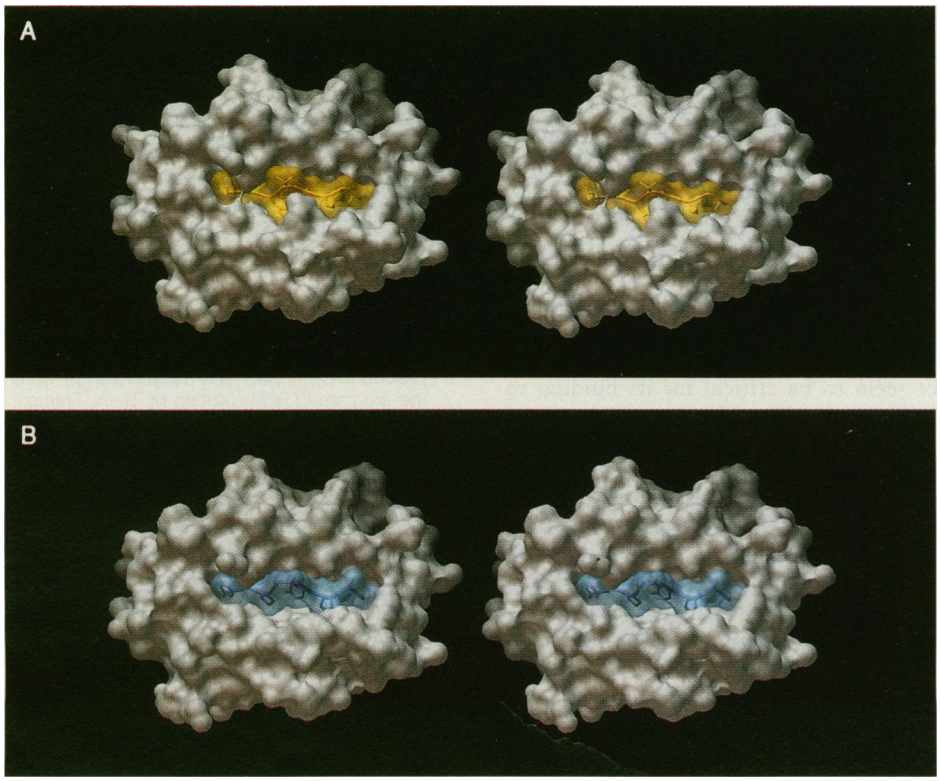

Fig. 6. Peptide binding site of H-$2K^b$. The solvent accessible surfaces (*51*) of the VSV-8 (**A**) and SEV-9 (**B**) peptides complexed with H-$2K^b$ are shown as stereo pair representations. Surfaces of the VSV-8 (yellow) and SEV-9 (cyan) are shown transparently. Peptides are displayed as tubes. Surfaces of the $\alpha_1\alpha_2$ domain (solid) were calculated separately from the peptide and illustrate that space is available in the groove for bound water molecules [close-up view in figure 7 (*22*)]. The most conspicuous difference between the two H-$2K^b$ surfaces is in the middle of the lower edge of the binding groove, corresponding to the position of Arg^{155}, which partially covers the peptide in VSV-8.

signal transduction. Whether any such role for segmental flexibility exists for MHC class I molecules is unknown, although β_2m plays a major role in the stabilization of the peptide-assembled complex (*16, 26*).

Both peptides bind to H-2K^b in extended conformations. The same observation has been made for the mixture of peptides associated with HLA-B27 (*11*). The extended conformation of bound peptides allows extensive hydrogen bonding with the peptide main chain. Helical conformations, by contrast, cannot undergo the same number of favorable intermolecular interactions because of their intramolecular backbone hydrogen bonds (*33*). Thus, the high affinity of peptide binding to MHC class I is in good measure the result of polar interactions with the extended peptide main-chain, which places very few restrictions on the functional groups at given side chain positions. This binding scheme, which may well be general for all MHC class I molecules, achieves high binding affinity to a diverse set of peptide sequences. One might predict that these same principles would hold true for MHC class II–peptide interactions (*34*).

In the SEV-9 peptide complex structure, the insertion of an additional residue beyond the optimal octamer length is accommodated not by extending the peptide at the NH_2- or COOH-terminus but by generating a central bulge in the peptide such that it kinks out of the groove. As to the binding of even longer peptides to H-2K^b and other class I molecules, it would appear from the extensive polar interactions at the peptide termini that additional residues extending in either direction would be energetically unfavorable, especially at the NH_2-terminus. In fact, the peptide NH_2-terminus is deeply embedded in the H-2K^b binding groove, whereas the COOH-terminus is held by a salt bridge near the surface. Consistent with these structural results are the observations that for the VSV-8 peptide, a two-residue COOH-terminal extension only decreases the affinity by a factor of 6, whereas the same extension at the NH_2-terminus reduces the affinity by ~100 (*17*).

That the MHC class I molecules are so effective in binding a diverse but restricted set of peptides with high affinity can be explained by the structure. The groove is fabricated in such a way that only side chains of MHC interact with peptide. The β-sheet platform forms a rigid framework on which the two helical walls can make slight adjustments, such as those seen for α_2 H1. That helices form the walls is advantageous because they can undergo segmental movements without disrupting the overall structure of the peptide binding groove. The segmentation of the helices in the α_2 domain further reinforces this notion. Another feature is that MHC class I side chains are free to adopt altered torsional angles providing optimal interaction with peptide. Further enhancing the flexibility of peptide binding are the bound water molecules, which move freely to bridge hydrogen bonds between peptide and MHC class I residues.

With respect to these structures, some observations should be made concerning TCR recognition. The resemblance of the TCR to the Fab fragment of an antibody (*35, 36*) gives rise to the expectation that the surface of MHC class I recognized by the TCR would be of a size similar to that of a protein antigen [700–900 Å^2 of surface made up by 14 to 21 residues (*37*)]. However, we show here that the bound peptides are extensively buried into the groove formed between the α_1 and α_2 helices (Fig. 6, A and B). Only a few of the peptide side-chains are significantly exposed after binding to H-2K^b, and they may contribute only (in an ideal association) some 100 to 300 Å^2 of surface to the TCR–MHC class I interface. Thus, if the analogy with antibodies holds true, at least two-thirds of what the TCR must recognize is MHC class I, not peptide.

Another issue is whether peptide associated conformational changes in the MHC class I molecule can affect TCR recognition. We have presented evidence here that both side chain and main chain variation occurs in H-2K^b as a consequence of the binding of two different peptides. The most profound peptide associated conformational changes occur around the upper surface of the peptide binding groove, the region most likely encountered by the TCR. The single substitution of an exposed peptide residue is sufficient to abrogate TCR recognition (*38*), although changes in the peptide do not seem to be critical for its binding to MHC class I (*39*). This is reminiscent of the observation that a single substitution in lysozyme (Gln^{121} → His) completely abolishes binding of antibody D1.3 (*40*). Similarly, even minor alterations in the MHC class I conformation in some cases may be sufficient to prevent association with a TCR. This is consistent with serological epitope mapping of peptide-loaded MHC class I molecules (*41*). The inference is that the interaction of the TCR with MHC class I molecules must be exquisitely tailored in order to recognize peptide sequence differences either directly or indirectly. Indeed, the MHC class I conformational changes identified here can be considered as synergistic effects, whereby the information communicated to the TCR by an individual peptide is augmented through its interaction with MHC class I molecules.

REFERENCES AND NOTES

1. P. Marrack and J. Kappler, *Science* **238**, 1073 (1987).
2. A. Townsend and H. Bodmer, *Annu. Rev. Immunol.* **7**, 601 (1989).
3. P. Kourilsky and J.-M. Claverie, *Adv. Immunol.* **45**, 107 (1989).
4. P. J. Bjorkman and P. Parham, *Annu. Rev. Biochem.* **59**, 253 (1990).
5. T. J. Tsomides and H. Eisen, *J. Biol. Chem.* **266**, 3357 (1991).
6. K. Matsui *et al.*, *Science* **254**, 1788 (1991).
7. P. J. Bjorkman *et al.*, *Nature* **329**, 506 (1987).
8. ———, *ibid.*, p. 512.
9. M. A. Saper, P. J. Bjorkman, D. C. Wiley, *J. Mol. Biol.* **219**, 277 (1991).
10. T. P. J. Garrett, M. A. Saper, P. J. Bjorkman, J. L. Strominger, D. C. Wiley, *Nature* **342**, 692 (1989).
11. D. R. Madden, J. C. Gorga, J. L. Strominger, D. C. Wiley, *ibid.* **353**, 321 (1991); *Cell*, in press.
12. B. P. Chen and P. Parham, *Nature* **337**, 743 (1989).
13. K. Falk, O. Rotzschke, S. Stevanovic, G. Jung, H.-G. Rammensee, *ibid.* **351**, 290 (1991).
14. T. S. Jardetzky, W. S. Lane, R. A. Robinson, D. R. Madden, D. C. Wiley, *ibid.* **353**, 326 (1991).
15. D. F. Hunt *et al.*, *Science* **255**, 1261 (1992).
16. M. R. Jackson, E. S. Song, Y. Yang, P. A. Peterson, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
17. M. Matsumura, Y. Saito, M. R. Jackson, E. S. Song, P. A. Peterson, *J. Biol. Chem.*, in press.
18. E. A. Stura *et al.*, *J. Mol. Biol.*, in press.
19. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.
20. G. M. Van Bleek and S. G. Nathenson, *Nature* **348**, 213 (1990).
21. T. N. M. Schumacher *et al.*, *ibid.* **350**, 703 (1991); W. M. Kast *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2283 (1991).
22. M. Matsumura, D. H. Fremont, P. A. Peterson, I. A. Wilson, *Science* **257**, 927 (1992).
23. The following programs were used for structural analysis: superimposition and distance measurements, INSIGHTII (Biosym Corporation, San Diego, CA); secondary structure definitions, DSSP [W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983)]; van der Waals contacts and hydrogen bonds, CONTACTSYM [S. Sheriff, W. A. Hendrickson, J. L. Smith, *J. Mol. Biol.* **197**, 273 (1989)].
24. D. J. Leahy, R. Axel, W. A. Hendrickson, *Cell* **68**, 1145 (1992).
25. To superimpose individual domains we used the program OVRLAP [M. G. Rossmann and P. Argos, *J. Biol. Chem.* **250**, 7525 (1975)]. To calculate pseudo rotation axes and translations we used the program ELBOW (written by D. H. Fremont and T. O. Yeates).
26. A. Townsend *et al.*, *Cell* **62**, 285 (1990).
27. A. Vitiello, T. A. Potter, L. A. Sherman, *Science* **250**, 1423 (1990).
28. S. Kozlowski *et al.*, *Nature* **349**, 74 (1991).
29. To calculate buried surfaces we used the program MS [M. L. Connolly, *J. Appl. Crystallogr.* **16**, 548 (1983)] and a 1.4 Å probe radius.
30. The values reported in the text correspond to the surface buried to a 1.4 Å probe. For comparison with other authors, the peptide areas buried to a 1.7 Å probe are 708 Å^2 (87 percent) for VSV-8 and 634 Å^2 (79 percent) for SEV-9. The corresponding buried areas for the H-2K^b binding groove are 921 and 856 Å^2, respectively. Bound water molecules were not included in these calculations. Similar length peptides (heptamers and octamers in antibody complexes have smaller buried surface areas. In Fabs B13I2 [R. L. Stanfield, T. M. Fieser, R. A. Lerner, I. A. Wilson, *Science* **248**, 712 (1990)] and 17/9 [J. M. Rini *et al.*, *ibid.* **255**, 959 (1992)] the corresponding buried surface areas are 490 Å^2 (66 percent) and 420 Å^2 (59 percent) for the peptide and 540 Å^2 and 470 Å^2 for the Fab, respectively.
31. D. H. Fremont, E. A. Stura, M. Matsumura, P. A. Peterson, I. A. Wilson, unpublished results.

32. I. A. Wilson, J. M. Rini, D. H. Fremont, G. G. Fieser, E. A. Stura, *Methods Enzymol.* **203**, 153 (1991).
33. D. Rognan, M. J. Reddehase, U. H. Koszinowski, G. Folkers, *Proteins Struct. Funct. Genet.* **13**, 70 (1992).
34. J. H. Brown *et al.*, *Nature* **332**, 845 (1988).
35. M. M. Davis and P. J. Bjorkman, *ibid.* **334**, 395 (1988).
36. C. Chothia, D. R. Boswell, A. M. Lesk, *EMBO J.* **7**, 3745 (1988).
37. D. R. Davies, E. A. Padlan, S. Sheriff, *Annu. Rev. Biochem.* **B59**, 439 (1990).
38. K. Shibata, M. Imarai, G. M. van Bleek, S. Joyce, S. G. Nathenson, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3135 (1992).
39. T. S. Jardetzky *et al.*, *EMBO J.* **9**, 1797 (1990); J. L. Maryanski, A. S. Verdini, P. C. Weber, R. R. Salemme, G. Corradin, *Cell* **60**, 63 (1990).
40. A. G. Amit, R. A. Mariuzza, S. E. V. Phillips, R. J. Poljak, *Science* **233**, 747 (1986).
41. J. A Bluestone, S. Jameson, S. Millár, R. Dick, in preparation; B. Catipovic, J. DalPorto, M. Mage, T. E. Johansen, J. P. Schneck, in preparation.
42. F. R. Carbone and M. J. Bevan, *J. Exp. Med.* **169**, 603 (1989).
43. A. J. Howard *et al.*, *J. Appl. Crystallogr.* **20**, 383 (1987).
44. P. D. M. Fitzgerald, *ibid.* **21**, 273 (1988).
45. The program HARADA (written by D. Filman and J. Arevalo) is based on a previously published algorithm [Y. Harada, A. Lifchitz, J. Berthou, *Acta Crystallogr.* **A37**, 398 (1981)].
46. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987).
47. T. A. Jones, *J. Appl. Crystallogr.* **11**, 268 (1978).
48. V. Luzatti, *Acta Crystallogr.* **5**, 802 (1952).
49. M. Carson, *J. Mol. Graphics* **5**, 103 (1987).
50. Tubes were generated with the program MCS TUBES (written by Y. Chen and A. Olson) and were rendered in the AVS environment [C. Upson *et al.*, *IEEE Comput. Graph. Appl.* **9**(no. 4), 30 (1989)].
51. To calculate solvent accessible surfaces we used a 1.4 Å radius probe and the program PQMS (written by M. L. Connolly), which is similar to the program MCS [M. L. Connolly, *J. Mol. Graph.* **3**, 19 (1985)].
52. We thank M. Pique for preparation of the figures, Y. Saito for production of the soluble H-2K^b, and J. Arevalo, N. Klinman, J. Bluestone, and R. Stanfield for helpful discussions. Supported in part by NIH grant CA-97489 (to P.A.P.) and training grant CA-09523 (to D.H.F.). This is publication 7467-MB from the Scripps Research Institute. Coordinates have been deposited in the Brookhaven Protein Data Bank and are available for prerelease.

24 June 1992; accepted 21 July 1992

Emerging Principles for the Recognition of Peptide Antigens by MHC Class I Molecules

Masazumi Matsumura, Daved H. Fremont,* Per A. Peterson, Ian A. Wilson

Class I major histocompatibility complex (MHC) molecules interact with self and foreign peptides of diverse amino acid sequences yet exhibit distinct allele-specific selectivity for peptide binding. The structures of the peptide-binding specificity pockets (subsites) in the groove of murine H-2K^b as well as human histocompatibility antigen class I molecules have been analyzed. Deep but highly conserved pockets at each end of the groove bind the amino and carboxyl termini of peptide through extensive hydrogen bonding and, hence, dictate the orientation of peptide binding. A deep polymorphic pocket in the middle of the groove provides the chemical and structural complementarity for one of the peptide's anchor residues, thereby playing a major role in allele-specific peptide binding. Although one or two shallow pockets in the groove may also interact with specific peptide side chains, their role in the selection of peptide is minor. Thus, usage of a limited number of both deep and shallow pockets in multiple combinations appears to allow the binding of a broad range of peptides. This binding occurs with high affinity, primarily because of extensive interactions with the peptide backbone and the conserved hydrogen bonding network at both termini of the peptide. Interactions between the anchor residue (or residues) and the corresponding allele-specific pocket provide sufficient extra binding affinity not only to enhance specificity but also to endure the presentation of the peptide at the cell surface for recognition by T cells.

The human and murine class I major histocompatibility complex (MHC) molecules are each separable into three allelic series of peptide-binding cell surface proteins, whose obligatory function is to present peptide antigens to cytotoxic T cells (*1–4*). Because of the extensive genetic polymorphism of these class I molecules, most individuals in an outbred population express two allelic forms of each isotype. Thus, the total number of class I molecules in an individual is very small compared with the entire repertoire of T cell receptors that can recognize a vast array of such peptides in the context of class I molecules.

Iso- and allotypic variants of class I molecules differ in those residues that form the peptide binding groove (*5*, *6*). This observation suggests that individual class I molecules bind peptides with varying specificities. However, each class I molecule must be able to bind large sets of peptides to explain the broad discriminatory capacity of the immune system. Sequence analyses of peptides isolated from individual class I molecules have confirmed this notion (*7–10*). That is, endogenously bound peptides are restricted in length to eight or nine amino acids, and, despite the size limitation, almost any position in the peptide can accommodate many different amino acid residues. However, in certain positions, one or a few amino acid residues recur with considerably increased frequencies. Such residues, called "anchor" residues, are characteristic for each class I allotype examined, and presumably the side chains of these residues interact with allotype-specific residues in the peptide binding groove. However, in the absence of high-resolution crystallographic data on the class I molecule bound to defined peptide sequences, there has been little direct evidence to support this hypothesis.

Two structures of murine H-2K^b molecules complexed with different viral peptides have now been elucidated at 2.3 and 2.5 Å resolution (*11*). This information, in conjunction with the available information on the structures of three human histocompatibility antigen (HLA) class I molecules (*6*, *12–14*) and the allele-specific peptide motifs (*9*), has permitted us to examine how class I molecules recognize and present a variety of chemically distinct peptides. We now show that conserved pockets at both ends of the peptide-binding groove accommodate the NH_2- and COOH-termini of peptides, whereas a deep polymorphic pocket in the middle of the groove forms the structural complementarity to the main anchor residue of the peptide and, hence, determines allele-specific peptide binding. Such limited usage of pockets appears to explain the broad but specific binding of peptides to MHC class I molecules.

Overall structure of peptide binding groove. We investigated the molecular mechanisms for MHC class I–peptide recognition by comparing structures of the murine class I molecule H-2K^b complexed with either VSV-8 peptide [vesicular stomatitis virus nucleoprotein(52–59); RGYVYQGL (*15*)] (*7*) or SEV-9 peptide [Sendai virus nucleoprotein(324–332); FAPGNYPAL) (*16*)] with the structure of the human class I molecule HLA-A2 (*5*,

M. Matsumura and P. A. Peterson are in the Department of Immunology and D. H. Fremont and I. A. Wilson are in the Department of Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037.

*D.H.F. is a graduate student in the Department of Chemistry, University of California–San Diego, La Jolla, CA 92093.