Supporting Information

Theodossis et al. 10.1073/pnas.1000032107

Database Mining

Consideration of MHC-I Restricted Peptides Involved in Crystal Contacts. In 101 of the binary complexes the peptides were found to be participating in symmetry related (i.e., crystal) contacts. The effect of such contacts on peptide conformation is a matter of debate, although in our experience it can vary in severity from negligible to moderate, depending on the extent of the contacts and the groups that are involved (e.g., main chain vs. side chain).

The PDB provides highly illustrative examples of the range of crystal contact-mediated effects on peptide conformation in the form of three binary pMHC-I structures 1I7R, 1HHK, and 3BEW. The first two are HLA-A*0201/9-mer structures, each consisting of two pMHC copies per ASU. In both structures the peptide environment varies slightly between the two copies. Whereas in one copy the peptide is not involved in crystal contacts, in the second a peptide side chain forms a single contact with a symmetry-related molecule. In both structures, the two copies adopt near identical conformations with an rmsd for all peptide atoms of <0.19 Å. A similar change in peptide environment between the two copies of the BF2*2101/10-mer complex in entry 3BEW, on the other hand, is accompanied by significant side-chain differences (rmsd of 1.2 Å for all peptide atoms). The most striking example, however, is that of a HLA-Cw*0401/9-mer binary complex (1QQD) and its KIR2DL1 ternary counterpart (1IM9). The peptide adopts equivalent conformations in the binary and ternary complex (rmsd of 0.33 Å and 0.67 Å for main chain and all peptide atoms, respectively), although some significant side-chain shifts are observed. The 1IM9 structure, however, also consists of a second pMHC-I copy in the ASU that is not part of a ternary complex. The peptide in this molecule forms extensive crystal contacts with the α-3 domain of a symmetry-related pMHC-I and as a result displays significant shifts along its backbone with respect to the peptide conformation observed in the binary structure (main chain and all atoms rms deviations of 0.84 Å and 1.26 Å, respectively).

In addition to the large conformational changes described in the cases above, it should be generally accepted that crystal contacts can often lead to otherwise flexible, solvent exposed peptide side chains, appearing in a single, well-ordered conformation. As such groups were to be the main targets of this analysis, we chose to eliminate from our sample population all binary pMHC-I entries displaying peptide-mediated crystal contacts (≤4.0 Å). In the case of structures consisting of more than one pMHC molecule in the ASU, only peptide copies not involved in crystal contacts were analyzed.

Impact of MHC-I Allelic Variation on the "Constrained" Nature of Bound Peptides. The structure of the 11-mer EBV peptide HPVGEADYFEY has been solved in complex with B*3501 (2FYY), in which it appears partially disordered, as well as B*3508 (2FZ3), in which it is well ordered and displays "type III" constraints. These differences in peptide conformation can be accounted for fully in terms of the HLA polymorphism at residue 156, which in the B*3508 structure has the effect of stabilizing one of the two main flanking residues (P5-Glu) that constrain the central bulge of the peptide.

Two other peptides, the synthetic 9-mer GRFAAAIAK and the EBV-derived 10-mer APQPAPENAY, were assigned to the type II constraints category based on their conformation when in complex with HLA-B*2705 (1JGE) and HLA-B*3501 (2AXG), respectively. Yet, these same peptides were found to display type III constraints in the context of HLA-B*2709 (1K5N) and HLA-B*3508 (2AXF), respectively. Moreover, the HSV glycoprotein B-derived 8-mer SSIEFARL was found displaying type III constraints in the context of H2Kb (1T0M), but was unconstrained in the context of $H2Kb^{m8}$ (1RJY). In all cases, the differences observed in the conformation and constraints displayed by the peptides can be accounted for by the HLA polymorphisms.

These four cases of MHC-dependent changes in peptide constraints account for 33% of the peptide sequences that were analyzed in the context of two or more MHC alleles (12 peptides in total) and provide clear evidence of the overriding importance that MHC sequence can have on peptide conformation in certain cases.

Role of pMHC-I Epitope Constraints on TCR Recognition and Binding. Given the small number of examples available for analysis, the role of peptide "self-constraints" in the interaction between MHC-bound peptide and TCR cannot be determined. Cases were identified in our analysis where constrained peptides either retained their binary complex conformation upon ligation or underwent conformational change (see below). It is worth noting, however, that where the conformation of a constrained peptide was conserved between its binary and ternary structures, the constraining interactions observed in the binary pMHC complex were often lost, indicating than the conformation of these peptides "relaxes" upon TCR ligation. These observations suggest that constraints may only serve to promote initial recognition by the cognate TCR, while subsequently allowing conformational changes to take place that are necessary for complete docking of the pMHC to the TCR.

Type I. The HLA-A*0201 restricted p1049 peptide (ALWGFFPVL; 1B0G), displays type I constraints, as the prominently surface-exposed side chain of P5-Phe is constrained by P3-Trp and P7-Pro (Fig. 2*B*). Interestingly, p1049 maintains its binary complex conformation upon ligation with the AHIII TcR (1LP9), with an rmsd between the two structures <0.36 Å for all peptide atoms, although binding to the TCR is accompanied by a partial loss of intrapeptide constraints. In three additional ternary complexes between AHIII and p1049 bound to HLA-A*0201 variants containing alanine substitutions in the Ag-binding cleft—K66A (2J8U), T163A (2UWE), and W167A (2JCC)—the peptide conformation is again conserved with respect to its binary structure (all atom rmsd <0.5 Å). Moreover, in the ternary complex of the K66A mutant (2J8U), the type I constraints of p1049 are also maintained. The HLA-A*0201 restricted GILGFVFTL peptide derived from influenza matrix protein, also displays type I constraints, as the conformation of the P5-Phe side chain is constrained by P3-Leu and P7-Phe (2VLL). $HLA\text{-}A*0201^{GILGFVFTL}$ has also been solved in ternary complex with the JM22 TCR (1OGA, 2VLJ, and 2VLK), and JM22-S99A mutant (2VLR), at different resolutions and in different space groups. The conformation of the peptide observed in the binary complex is also conserved in each of the ternary complexes (rmsd < 0.37 Å for all peptide atoms), whereas the type I constraints are also maintained in the two higher resolution ternary complexes (1OGA and 2VLJ).

Type II. In the case of the HTLV-1 TAX derived peptide LLFGYPVYV bound to HLA*0201 (2AV7), binding to either the B7 (1BD2) or A6 (1AO7) TCR is accompanied by large conformational rearrangements centered on P6-Pro (all atoms rmsd >1.59 Å). Despite these movements, P6-Pro and P3-Phe are able to maintain the type II constraints on P5-Tyr in the ternary complexes.

Type III. The conformation adopted by the 13-residue (LPEPLPQGQLTAY) determinant from EBV when bound to HLA-B*3508 alone is also maintained in its ternary complex with the SB27 TcR (2AK4), with an average rmsd for all peptide atoms of 0.75 Å, although the interaction between P6-10 and P10-Leu is not conserved. The conformation of the $H2K^b$ restricted 8-mer

KVITFIDL (pKB1 peptide; 1KJ3) also remains relatively fixed upon ligation by the KB5-C20 TcR (all atom rmsd <0.58 Å), whereas the $H2K^{bm8}$ restricted endogenous pBM8 peptide (2CLV; SQYYYNSL) undergoes a significant conformational change upon binding to the BM3.3 TCR (2OL3; all atom rmsd of 1.3 Å), although without loss of constraints.

Correlation Between Constraint and Immunogenicity. By characterizing antigenic peptides as being unconstrained in their pMHC-bound states, no assumptions are made about their relative immunogenicity, which varies significantly between peptides assigned to this category. The HLA-B*3501-restricted 11-mer EPLPQGQLTAY peptide derived from the BZLF1 antigen of Epstein-Barr virus displays an unconstrained conformation in the binary structure (1ZSD). Nevertheless, the HLA-B*3501/EPL complex generates a significant CTL response in EBV-infected individuals, even though the structure of its ternary complex with the ELS4 TCR (2NX5) has shown that the peptide adopts a significantly altered conformation upon TCR docking. The unconstrained HLA-B*2705 restricted RRIYDLIEL peptide derived from EBV is also reported as being immunodominant (2BSR), as is the HLA-A*0201 restricted 9-mer (AAGIGILTV) derived from MART-1 (2GUO; Fig. 3*F*). By contrast, the A2L mutant of that MART-1 peptide (2GTZ), which is also unconstrained, displays a greatly reduced, or abolished CTL response.

Our analysis of published pMHC-I structures led to the identification of a number of instances where alterations in peptide constraints, and subsequent changes in pMHC-I topology, correlated with loss of pMHC-I antigenicity. However, the introduction of constraints in a peptide (and the increase in peptide rigidity implied by this) does not necessarily correlate with greater immunogenicity. The B*3501 restricted 14-mer LPAVVGLSPGEQEY derived from an alternative reading frame of M-CSF (1XH3), displays a helical conformation between positions 9 and 12. In addition to the characteristic helical main chain hydrogen bonding network present within this region, the peptide is stabilized by multiple vdw and hydrogen bonding interactions involving main chain and side chain groups between positions P5-Val and P13-Glu. Despite the large number of constraining interactions, this peptide is marked by considerable flexibility in its central region. By contrast, two variants of this peptide containing a single alanine substitution at position P2-Pro and a double alanine substitution at P2-Pro and P9-Pro, were shown to display significant positional differences between positions P4 and P12 with respect to the index peptide, although the two maintain a helical conformation between P9 and P12. Moreover, introduction of each mutation was correlated with a decrease in flexibility with respect to the index peptide, which in turn was accompanied by a decrease in immunogenicity. Thus, the P2A mutant, displaying intermediate flexibility, acted as a weak agonist, whereas the P2A/P9A, being the most rigid of the three peptides, completely abrogated the T-cell response. These findings demonstrate that single amino acid changes in a highly constrained peptide can have significant effects on its MHC-bound conformation and that decreased peptide flexibility is not necessarily consistent with greater immunogenicity.

Structure of $H2D^b$ in Complex with PA_{224}-F6A. The crystal structure of H-$2D^b$ PA-F6A was solved in space group $P4_32_12$ and consists of one molecule in the asymmetric unit. Clear electron density for the backbone and side chains of the MHC molecule was observed throughout the modeled structure. The SSLENARAYV peptide adopts an extended conformation and forms anchoring interactions with the α1-α2 domains at the peptide positions P2-Ser (Tyr7, Tyr45, and Glu163 of the H-$2D^b$ heavy chain), P5-Asn (Gln70, Trp73, and Gln97), and P10-Val (Ser77, Leu81, and Trp147). Additionally there are vdw contacts between the MHC molecule and the side chain of P1-Ser (Lys68 and Glu163), and the backbone at P4-Glu (His155), P6-Ala (Tyr156 and Trp173), and P9-Tyr (Lys146 and Trp147). The region of the peptide between P6-Ala and P9-Tyr extends prominently from the peptide-binding groove and is significantly solvent exposed. This includes the site of the P6-Phe to P6-Ala mutation and the neighboring P7-Arg.

To directly compare the crystal structure of the F6A mutant to the wild-type $H2D^bPA$ structure, some investigation was carried out regarding the differences in the peptide conformation between the two published structures available in the PDB (1YN6 and 1WBY). In 1WBY, the prominent P7-Arg in the peptide forms a salt-bridging interaction with P4-Glu of the peptide. This means that the side chain of the arginine has moved away from P6-Phe where it has a significant interaction surface in the alternative structure, 1YN6. Upon downloading electron density maps from the Uppsala EDS, the picture was not made any clearer as the P7-Arg side chain had broken density past the Cγ in 1WBY. Furthermore there were no maps available for 1YN6. As the structure of $H2D^bPA$ was solved in this laboratory, a refinement could be carried out using archived data files. This also ensured that an identical strategy to the mutant F6A structure could be used to provide a fair comparison.

From the archived $H2D^bPA$ data files, exactly the same steps were carried out as for the F6A mutant—data processing with HKL2000, molecular replacement by PHASER, model building in COOT 0.5, and restrained refinement followed by TLS refinement in REFMAC5. The final statistics for this refinement were R_{work} = 19.2% and R_{free} = 25.2% (Table S1). This is comparable to the published structure, and the validation statistics from MOLPROBITY are as good as those obtained for the F6A structure. Density maps clearly show that the P7-Arg of the peptide is in the conformation seen in 1YN6, forming vdw interactions with P6-Phe.

Comparing the F6A structure and the wild-type H2-D^bPA structure that has been refined in the same manner, there is little movement between the α1 and α2 domains of the MHC molecule (rmsd 0.3 Å over residues 1–180 of the α-chain, at the Cα position). The two peptides adopt equivalent conformations, overlaying with an rmsd of 0.16 Å at the Cα position, with identical anchoring interactions to the MHC molecules. The major difference between the two structures occurs at the solvent-exposed C-terminal region of the F6A mutant peptide. The absence of the phenylalanine side chain present in the wild-type epitope is compensated for by the presence of an additional water molecule and a slight movement in the side chain of Glu41 from the HC of a neighboring symmetry-related molecule. The side chain of P7-Arg is shifted significantly in the F6A structure compared to the wild type, where it forms a number of vdw interactions with the aromatic side chain of the neighboring phenylalanine. In the absence of these stabilizing interactions the arginine side chain has relatively shifted toward the N-terminal end of the peptide. In this new position it forms vdw interactions with the side chain of symmetry-related Glu41, as well as the main chain at Glu41 and Ala40 of the neighboring heavy chain. The magnitude of the movement of the side chain is such that the guandinium head group of P7-Arg has moved by 8.5Å, with the overlaid side chains forming an angle of 100° from the equivalent positions of the Cα atom.

Thermostability assay. The thermostability of recombinant MHC/peptide complexes was assessed using circular dichroism (CD). CD spectra were measured on a Jasco 810 spectropolarimeter using a thermostatically controlled cuvettete at temperatures between 35 °C and 70 °C as described in detail elsewhere (1). Far-UV spectra from 195 to 250 nm were collected and averaged over two individual scans; measurements for the thermal melting

experiments were made at intervals of 0.1 °C at a rate of 1 °C/min. The midpoint of thermal denaturation (Tm) for each protein was calculated by taking the first derivative of the ellipticity data and identifying the inflection point.

Structure Analysis and Figure Generation. All structures were analyzed using COOT version 0.6-pre. Coordinate files were edited using the CCP4i implementation of PDBSET. Root mean square deviation (rmsd) calculations and molecular contact analysis were carried out using LSQKAB and CONTACT, respectively. Molecular contacts were further analyzed and visualized using the COOT implementation of REDUCE and PROBE. All figures were generated using PyMOL. In the case of *B–E* in Fig. 2, images were initially generated using the COOT implementation of RASTER3D and then imported into PyMOL.

1. Webb AI, et al. (2004) The structure of H-2K(b) and K(bm8) complexed to a herpes simplex virus determinant: Evidence for a conformational switch that governs T cell repertoire selection and viral resistance. *J Immunol* 173:402–409.

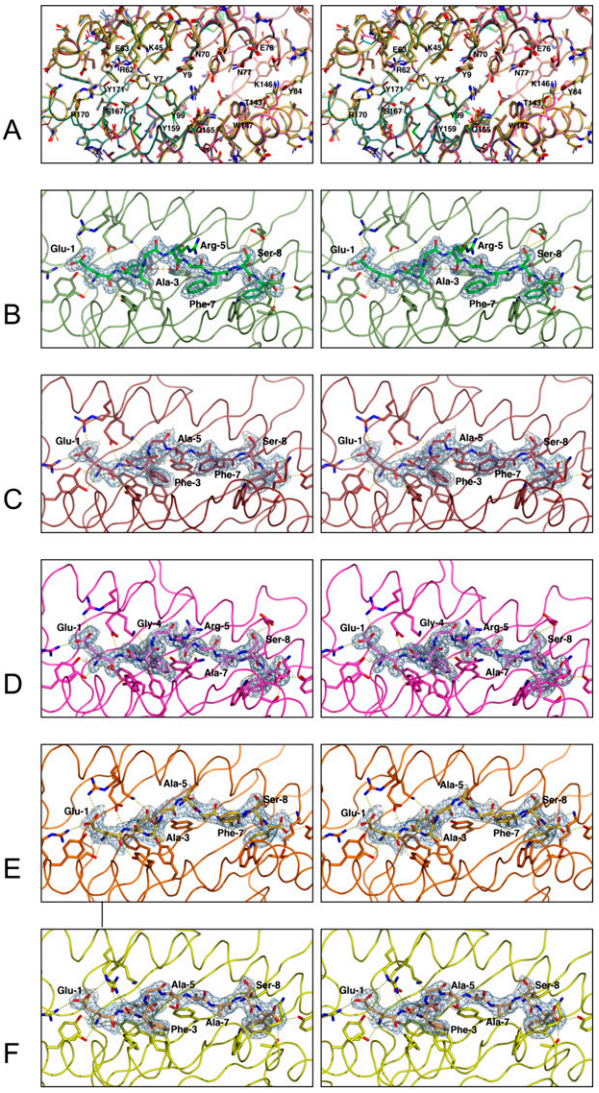


Fig. S1. Stereographic representations of the Ag-binding cleft topology and peptide electron density for the HLA-B*4402/DPα structure and each of the five APL complexes. (*A*) Superposition of the α1/α2 domains of HLA-B*4402 from the index DPα structure (1M6O; cyan) with each of the APL complexes (F3A, green; F5A, red; F7A, magenta; F3A/R5A, orange; and R5A/F7A, yellow) gives rmsd values of 0.36–0.61 Å for all atoms of residues 3–180. The heavy chain backbone is represented in ribbon format, with side chains shown as sticks. Residues that form significant electrostatic interactions with DPα are labeled. (*B*) Refined structure of the HLA-B*4402/DPα F3A complex. A ribbon representation is given of the heavy chain backbone (residues 1–181), whereas the bound peptide, as well as selected heavy chain side chains, are presented in stick format. Overlayed is the final *2Fo-Fc* electron density for the refined peptide contoured at 1.0 σ (gray mesh), as well as the unbiased *Fo-Fc* density within 1.8 Å of the peptide, contoured at +2.5 σ (marine mesh). Potential hydrogen bonds and salt bridges between the peptide and MHC are shown as yellow dashes. (*C–F*) Equivalent views of the F5A (red), F7A (magenta), F3A/R5A (orange), and R5A/F7A (yellow) complexes, respectively.

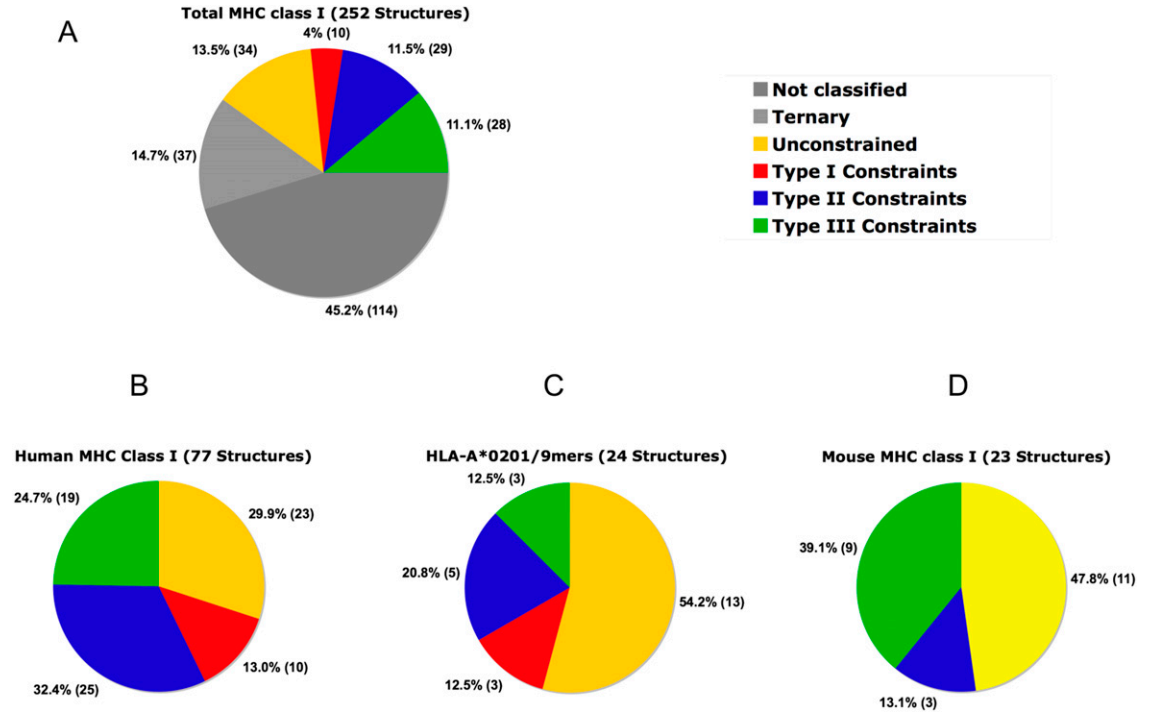


Fig. S2. (*A*) Assignment of the 252 MHC class-I entries in the PDB database according to peptide "self-constraints": (*i*) Ternary, structures of pMHC in complex with TCR or KIR. (*ii*) Not classified, (*a*) pMHC binary structures in which the peptide is involved in crystal contacts (≤4 Å); (*b*) three pMHC binary complex entries corresponding to a molecular model (1L6Q) and two crystal structures of MHC in complex with mixtures of endogenous peptides (1HSA and 1HSB); (*c*) 10 binary structures that have been duplicated at higher resolution and were therefore considered redundant. (*iii*) Unconstrained. (*iv*) Types I, II, and III constrained. (*B*) Distribution of peptide self-constraints among the 79 eligible binary HLA structures. (*C*) Distribution of self-constraints among the 25 eligible binary structures of HLA-A*0201 in complex with 9-mer peptides. (*D*) Distribution of self-constraints among the 28 eligible binary murine MHC structures.

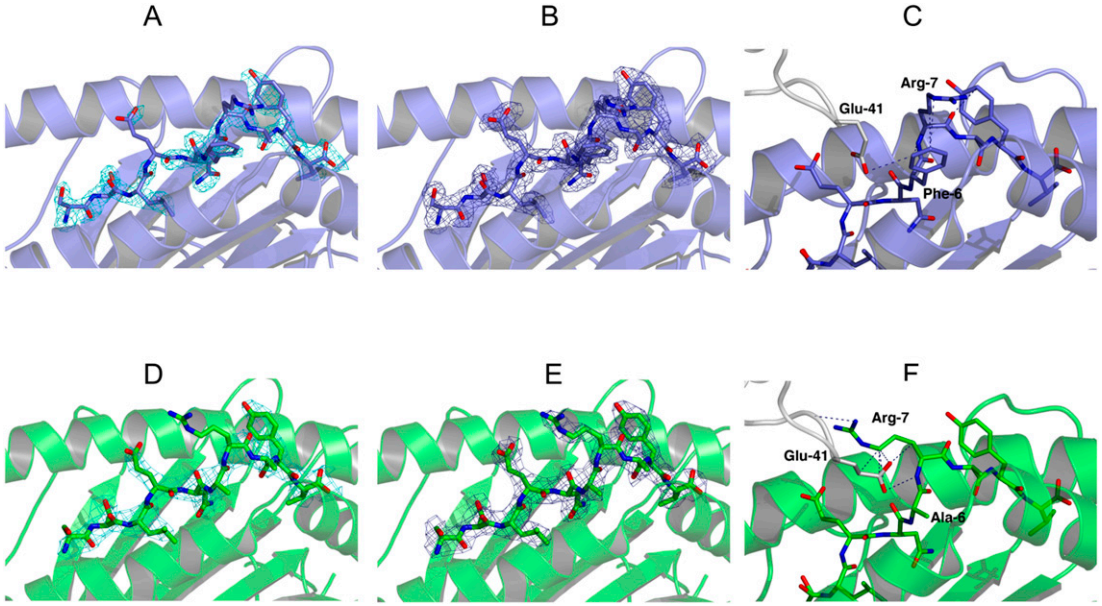


Fig. S3. The crystal structures of H2D^b in complex with PA_{224} and PA_{224}-F6A. (*A*–*C*) Refined coordinates for the H2D^b complex with the wild-type SSLENFRAYV peptide. (*A*) Unbiased *Fo-Fc* density, contoured at 2.5 σ, is colored cyan. (*B*) The final *2Fo-Fc* density, contoured at 1.0 σ, is colored blue. (*C*) A neighboring symmetry-related molecule that interacts with the peptide is shown in gray. Van der Waals interactions are shown as blue dashes. (*D* and *E*) Respective views for H2D^b/PA_{224}-F6A.

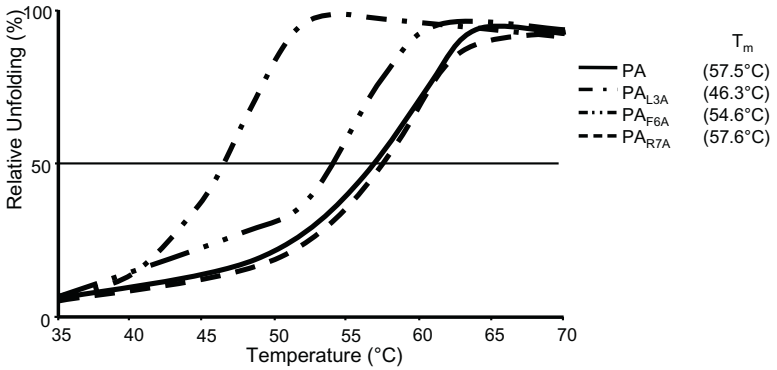


Fig. S4. H2D^b binding by PA mutants. The thermal stability of the WT PA_{224} and each variant peptide complexed with H2D^b was monitored by circular dichroism as the temperature was raised from 35 °C to 70 °C at 1 °C/min. CD spectra were measured as described in *Materials and Methods*. The curves represent the relative unfolding of protein.

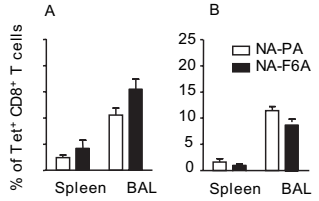


Fig. S5. Unaltered magnitude of acute D^bNP_{366} and $D^bPB1\text{-}F2_{62–70}$-specific responses after NA-PA or NA-F6A viral challenge. Lymphocytes were obtained as for Fig. 3 and stained for D^bNP_{366} (*A*) or $D^bPB1\text{-}F2_{62–70}$ (*B*) tetramers. All of the cells were subsequently stained with anti-CD8a-FITC antibodies. Shown is the percentage of Tetramer$^+$ CD8$^+$ T cells. Data are mean ± SD for groups of three mice.

Table S1. Data collection and refinement statistics for the structures of HLA-B* 4402 in complex with five DPα_{46} alanine substitution mutants and of H2-D^b in complex with PA$_{224}$ and its F6A mutant

	HLA-B*4402/ DPα_{46} APLs					H2-D^b/PA$_{224}$ and F6A mutant	
Data collection	EEAGRAFSF	EEFGAAFSF	EEFGRAASF	EEAGAASF	EEAGAASF	SSLENFRAYV	SSLENARAYV
Temperature (K)	100	100	100	100	100	100	100
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P4_32_12$	$P4_32_12$
Cell dimensions (*a*,*b*,*c*) (Å)	50.7,82.0,110.4	50.7,81.8,110.0	50.8,82.3,110.3	50.8,82.2,110.0	50.6,82.1,109.2	56.1,56.1,273.5	56.1,56.1,275.1
Resolution (Å)	1.8	2.1	1.7	2.3	2.5	2.2	2.7
Total observations	159,08 1	76,666	147,04 1	44,590	42,913	179,70 7	57,196
Unique observations	43,106	27,449	49,356	16,978	15,875	23,389	12,799
Average multiplicity	3.7	2.5	3.0	2.6	2.7	7.7	4.5
Data completeness(%)	99.1 (99.1)	96.6 (99.8)	94.2 (92.2)	84.8 (87.3)	93.6 (71.3)	99.6 (99.5)	98.2 (98.0)
$\langle I/\sigma_I \rangle$	7.0 (2.0)	5.5 (1.9)	20.4 (3.9)	5.3 (2.0)	10.6 (3.0)	43.4 (5.4)	18.4 (5.1)
R_{merge} (%)	7.6 (33.1)	7.7 24.4)	6.0 (28.6)	10.5 (27.5)	13.2 (24.5)	5.6 (20.8)	9.1 (31.7)
Refinement							
Resolution (Å)	65.9–1.8	34.0–2.1	65.9–1.7	41.0–2.4	65.6–2.6	32–2.2	28–2.7
R_{factor}(%)	18.8	20.5	19.7	23.3	21.1	19.2	22.0
R_{free}(%)	23.1	26.4	24.2	28.9	27.5	25.2	28.4
Nonhydrogen atoms							
Protein	320 1	316 7	318 6	315 5	312 1	316 6	316 0
Solvent	607	353	552	9 2	148	272	122
Rms deviations from ideal							
Bond lengths (Å)	0.013	0.01 0	0.016	0.011	0.005	0.014	0.00 6
Bond angles (°)	1.39	1.22	1.50	1.30	0.86	1.48 8	0.88 8
Ramachandran plot (%)							
Favored region	98.2	97.6	98.5	96.6	98.1	97.4	97.1
Allowed region	1.8	2.1	1.2	3.4	1.6	2.6	2.4
Disallowed	0.0	0.3	0.3	0.0	0.3	0.0	0.5
Average B-factors (Å^2)							
Main chain	19.3	26.0	19.6	38.4	22.8	28.7	26.4
Side chain	22.0	27.9	22.5	39.8	23.6	30.5	27.4
Solvent molecules	33.8	38.0	34.4	18.4	20.7	34.7	21.2

Values in parentheses are for the bin of highest resolution (approximate interval = 0.1 Å). $R_{merge} = \sum|I_{hkl} - \langle I_{hkl}\rangle|/\sum I_{hkl}$. $R_{factor} = \sum_{hkl} \big||F_o| - |F_c|\big|/\sum_{hkl}|F_o|$ for all data except ≈4–10%, which were used for R_{free} calculation.

Table S2. List of 101 MHC class I PDB entries by species and allele with self-constraint assignments

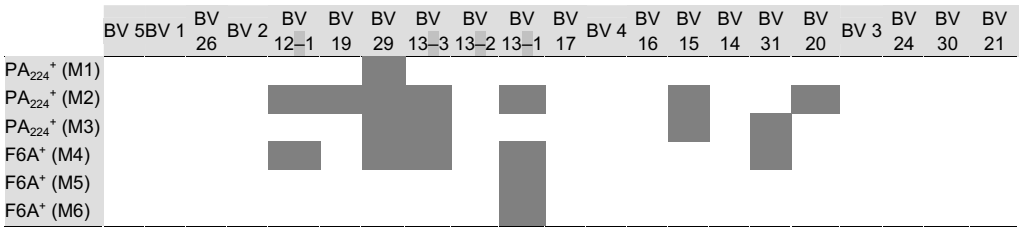
MHC	PDBid	Chains analyzed	Peptide sequence	Peptide length	Assigned category	Constraining interactions[†]
HLA-A*0201	1AKJ	C	ILKEPVHGV	9	Type II	K3⇒P5⇐H7
HLA-A*0201	1EEY	C, F	ILSALVGIV	9	Unconstrained	—
HLA-A*0201	1EEZ	C, F	ILSALVGIL	9	Type III	S3⇒A4⇐L5
HLA-A*0201	2VLL	C, F	GILGFVFTL	9	Type I	L3⇒F5⇐F7
HLA-A*0201	1I1F	C, F	FLKEPVHGV	9	Unconstrained	—
HLA-A*0201	1I1Y	C, F	YLKEPVHGV	9	Unconstrained	—
HLA-A*0201	2GTZ	C, F	ALGIGILTV	9	Unconstrained	—
HLA-A*0201	1QR1	C, F	IISAVVGIL	9	Type III	S3⇒A4⇐V5
HLA-A*0201	2GTW	C, F	LAGIGILTV	9	Type III	A2⇒G3,I4,G5⇐I6
HLA-A*0201	2GUO	C, F	AAGIGILTV	9	Unconstrained	—
HLA-A*0201	2V2W	C, F	SLYNTVATL	9	Type III	Y3⇒N4⇐T5,V6
HLA-A*0201	2V2X	C, F	SLFNTVATL	9	Unconstrained	—
HLA-A*0201	1T1W	C	SLFNTIAVL	9	Unconstrained	—
HLA-A*0201	1T1X	C	SLYLTVATL	9	Unconstrained	—
HLA-A*0201	1T1Z	C	ALYNTAAAL	9	Unconstrained	—
HLA-A*0201	1T20	C	SLYNTIATL	9	Unconstrained	—
HLA-A*0201	1TVB	C, F	ITDQVPFSV	9	Unconstrained	—
HLA-A*0201	1TVH	C, F	IMDQVPFSV	9	Unconstrained	—
HLA-A*0201	1B0G	C, F	ALWGFFPVL	9	Type I	W3⇒F5⇐P7
HLA-A*0201	1I7R	C	FAPGFFPYL	9	Unconstrained	—
HLA-A*0201	1I7T	C, F	ALWGVFPVL	9	Type I	W3⇒V5⇐P7
HLA-A*0201	1I7U	C, F	ALWGFVPVL	9	Type II	W3⇒F5⇐P7
HLA-A*0201[‡]	1HHK	C	LLFGYPVYV	9	Type II	F3⇒Y5⇐P6,V7
HLA-A*0201	2AV7	C	LLFGYPVYV	9	Type II	F3⇒Y5⇐P6,V7
HLA-A*0201	2CLR	C, F	MLLSVPLLLG	10	Type III	L3⇒S4⇐V5
HLA-A*0201	2GT9	C, F	EAAGIGILTV	10	Unconstrained	—
HLA-A1101	1X7Q	C	KTFPPTEPK	9	Type II	F3⇒P4⇐P5
HLA-B*0801	1AGD	C	GGKKKYKL	8	Type II	K3⇒K4⇐Y6
HLA-B*0801	1AGB	C	GGRKKYKL	8	Type II	R3⇒K4⇐Y6
HLA-B*0801	1AGC	C	GGKKKYQL	8	Type II	K3⇒K4⇐Y6
HLA-B*0801	1AGE	C	GGKKKYRL	8	Type II	K3⇒K4⇐Y6
HLA-B*0801	1AGF	C	GGKKRYKL	8	Type III	K3⇒K4⇐R5
HLA-B*1402	3BXN	C	IRAAPPPLF	9	Type II	A3⇒A4⇐A5
HLA-B*1501	1XR8	C	LEKARGSTY	9	Type II	K3, A4⇒R5
HLA-B*1501	1XR9	C	ILGPPGSVY	9	Unconstrained	—
HLA-B*1501	3C9N	C	VQQESSFVM	9	Type II	Q3⇒E4⇐S6
HLA-B*2705	1JGE	C	GRFAAAIAK	9	Type II	F3⇒A5⇐I7
HLA-B*2705	1W0V	C	RRLPIFSRL	9	Type I	L3⇒I5⇐S7
HLA-B*2705	2BSR	C	RRIYDLIEL	9	Unconstrained	—
HLA-B*2705	2BST	C	SRYWAIRTR	9	Type III	Y3⇒W4⇐A5
HLA-B*2705	3B6S	C	RRKWRRWHL	9	Type II	K3⇒CIR5⇐W7
HLA-B*2705	3BP4	C	IRAAPPPLF	9	Type II	A3⇒A4⇐P5
HLA-B*2705	2BSS	C	KRWIILGLNK	10	Type II	W3,I5⇒L6,G7⇐L8
HLA-B*2709	1K5N	C	GRFAAAIAK	9	Type III	F3⇒A4⇐A5
HLA-B*2709	1W0W	C	RRLPIFSRL	9	Type I	L3⇒I5⇐S7
HLA-B*2709	3B3I	C	RRKWRRWHL	9	Type II	K3,W4⇒R6⇐W7
HLA-B*2709	3BP7	C	IRAAPPPLF	9	Type II	A3⇒A4⇐P5
HLA-B*2709	1JGD	C	RRLLRGHNQY	10	Type III	L3⇒L4⇐R5⇒G6⇐N8
HLA-B*3501	1A9B	C, F	LPPLDITPY	9	Type II	D5⇒T7⇐P8
HLA-B*3501	2CIK	C	KPIVVLHGY	9	Unconstrained	—
HLA-B*3501	2AXG	C	APQPAPENAY	10	Type II	Q3⇒A5⇐P6
HLA-B*3501	2FYY	C	HPVGEADYFEY	11	Unconstrained	—
HLA-B*3501	1ZSD	C	EPLPQGQLTAY	11	Unconstrained	—
HLA-B*3501	1ZHK	C	LPEPLPQGQLTAY	13	Type III	P6⇒Q7,G8,Q9⇐L10
HLA-B*3501	1XH3	C	LPAVVGLSPGEQEY	14	Type III	V5,L7⇒S8,P9,G10⇐E11, Q12
HLA-B*3508	3BWA	C	FPTKDVAL	8	Unconstrained	—
HLA-B*3508	2AXF	C	APQPAPENAY	10	Type III	Q3,P4⇒A5,P6⇐E7
HLA-B*3508	2FZ3	C	HPVGEADYFEY	11	Type III	V3,E5⇒A6,D7,Y8⇐F9
HLA-B*3508	2NW3	C	EPLPQGQLTAY	11	Unconstrained	—
HLA-B*3508	3BW9	C	CPSQEPMSIYVY	12	Type II	S3,E5⇒P6,M7,S8,I9⇐Y10
HLA-B*3508	1ZHL	C	LPEPLPQGQLTAY	13	Type III	P6⇒Q7,G8,Q9⇐L10

Table S2. Cont.

MHC	PDBid	Chains analyzed	Peptide sequence	Peptide length	Assigned category	Constraining interactions[†]
HLA-B*4402	1M6O	C	EEFGRAFSF	9	Type I	F3⇒R5⇐F7
HLA-B*4403	1SYS	C	EEPTVIKKY	9	Unconstrained	—
HLA-B*4403	1N2R	C	EEFGRAFSF	9	Type I	F3⇒R5⇐F7
HLA-B*4405	1SYV	C	EEFGRAFSF	9	Type I	F3⇒R5⇐F7
HLA-B*5101	1E28	C	TAFTIPSI	8	Unconstrained	—
HLA-B*5101	1E27	C	LPPVVAKEI	9	Type II	P3⇒V4⇐V5,A6
HLA-B*5301	1A1N	C	VPLRPMTY	8	Type II	V1⇒R4⇐P5
HLA-B*5301	1A1M	C	TPYDINQML	9	Type I	Y3⇒I5⇐Q7
HLA-B*5301	1A1O	C	KPIVQYDNF	9	Type I	I3⇒Q5⇐D7
HLA-B*5701	2RFX	C	LSSPVTKSF	9	Type III	V5⇒T6⇐K7
HLA-B*5703	2BVQ	C	KAFSPEVI	8	Type II	F3⇒S4⇐P5
HLA-B*5703	2BVP	C	ISPRTLDAW	9	Type II	I1,S2,P3⇒R4⇐T5,L6
HLA-B*5703	2BVO	C	KAFSPEVIPMF	11	Type III	F3,P5⇒E6,V7⇐I8,P9
HLA-B*5703	2HJK	C	KGFNPEVIPMF	11	Type III	F3,P5⇒E6,V7⇐I8,P9
HLA-B*5703	2HJL	C	KAFNPEIIPMF	11	Type III	F3,P5⇒E6,I7⇐I8,P9
HLA-Cw*0401	1QQD	C	QYDDAVYKL	9	Type III	D3⇒D4⇐A5
$H2D^b$	2CII	C	YPAL	4	Unconstrained	—
$H2D^b$	1CE6	C	FAPGNYPAL	9	Unconstrained	—
$H2D^b$	1FFN	C, F	KAVYNFATM	9	Unconstrained	—
$H2D^b$	1FFO	C, F	AAVYNFATM	9	Unconstrained	—
$H2D^b$	1FFP	C, F	SAVYNFATM	9	Unconstrained	—
$H2D^b$	1QLF	C	FAPSNYPAL	9	Unconstrained	—
$H2D^b$	3BUY	C	LSLRNPILV	9	Unconstrained	—
$H2D^b$	3CC5	C, F	KVPRNQDWL	9	Unconstrained	—
$H2D^b$	3CCH	C, F, I, L	EGSRNQDWL	9	Type II	Q6⇒D7⇐W8
$H2D^b$	1YN7	C	SSLENFAAYV	10	Type III	F6⇒A7⇐A8
$H2K^b$	2ZSV	E, F	RAQIFANI	8	Type III	ABA2,Q3⇒I4⇐F5
$H2K^b$	2ZSW	M, N, O, P	RAYIFANI	8	Type III	ABA2,Y3⇒I4⇐F5
$H2K^b$	1VAC	P	SIINFEKL	8	Type II	I2,I3⇒N4⇐F5
$H2K^b$	2MHA	E, F	RGYVYQGL	8	Type III	Y3⇒V4⇐Y5
$H2K^b$	1NAN	M, Q	INFDFNTI	8	Type III	N2,F3⇒D4⇐F5
$H2K^b$	1KJ3	P, Q	KVITFIDL	8	Type III	V2,I3⇒T4⇐F5
$H2K^b$	1T0M	P, Q	SSIEFARL	8	Type III	I3⇒E4⇐F5
$H2K^b$	1WBZ	P, Q	SSYRRPVGI	9	Type II	Y3,R5⇒P6⇐V7G8,I9
$H2K^{bM8}$	1RJZ	P, Q	SEIEFARL	8	Unconstrained	—
$H2K^{bM8}$	2CLZ	C, M	INFDFNTI	8	Type III	N2⇒D4⇐F5
$H2K^{bM8}$	2CLV	C, M	SQYYYNSL	8	Type III	Q2,Y3⇒Y4⇐Y5
$H2K^{bM8}$	1RJY	P, Q	SSIEFARL	8	Unconstrained	—
$H2K^d$	1VGK*	C	SYVNTNMGL	9	Unconstrained	—
BF2*2101	3BEW	C	REVDEQLLSV	10	Type II	V3⇒D4,E5,Q6,L7⇐L8,S9

[†]Arrows point from flanking/constraining residues to the constrained central feature.
[‡]Mutation in heavy chain.

Table S3. Summary of the repertoire for different individuals showing an altered TRBV13-1 bias after NA-F6A PR8 virus infection



$CD8^+$ D^bPA_{224} or D^bPA-F6A cells were recovered from the spleen on d9 and sorted using a FACSAria and processed as described in *Materials and Methods*. M, individual mice tested; BV, IMGT nomenclature for TCRVβ gene segments.

Table S4. TCR CDR3β analysis within individual mice

CDR3β region		Jβ	Length	F6A$^+$ (M4) frequency	F6A$^+$ (M5) frequency	F6A$^+$ (M6) frequency
S G G G N T G Q L	Y F G	2S2	9	13/26		5/21
R G G G N T G Q L	Y F G	2S2	9	7/26		
K G G G N T G Q L	Y F G	2S2	9	4/26		
S D S T S N S D Y	T F G	1S2	9	2/26		
S N Q G D S D Y	T F G	1S2	8		13/26	
S D G D W G R A E T L	Y F G	2S3	11		12/26	
S E N W G S S Y E Q	Y F G	2S6	10		1/26	
S G G S N T G Q L	Y F G	2S2	9			6/21
S E S Y E Q	Y F G	2S6	6			5/21
S G G A N T G Q L	Y F G	2S2	9			3/21
S R T V N T G Q L	Y F G	2S2	9			2/21

The TRBV13-1 repertoire was analyzed from CD8+F6A+ cells sorted from spleen on day 9 after infection, as described in *Materials and Methods*.