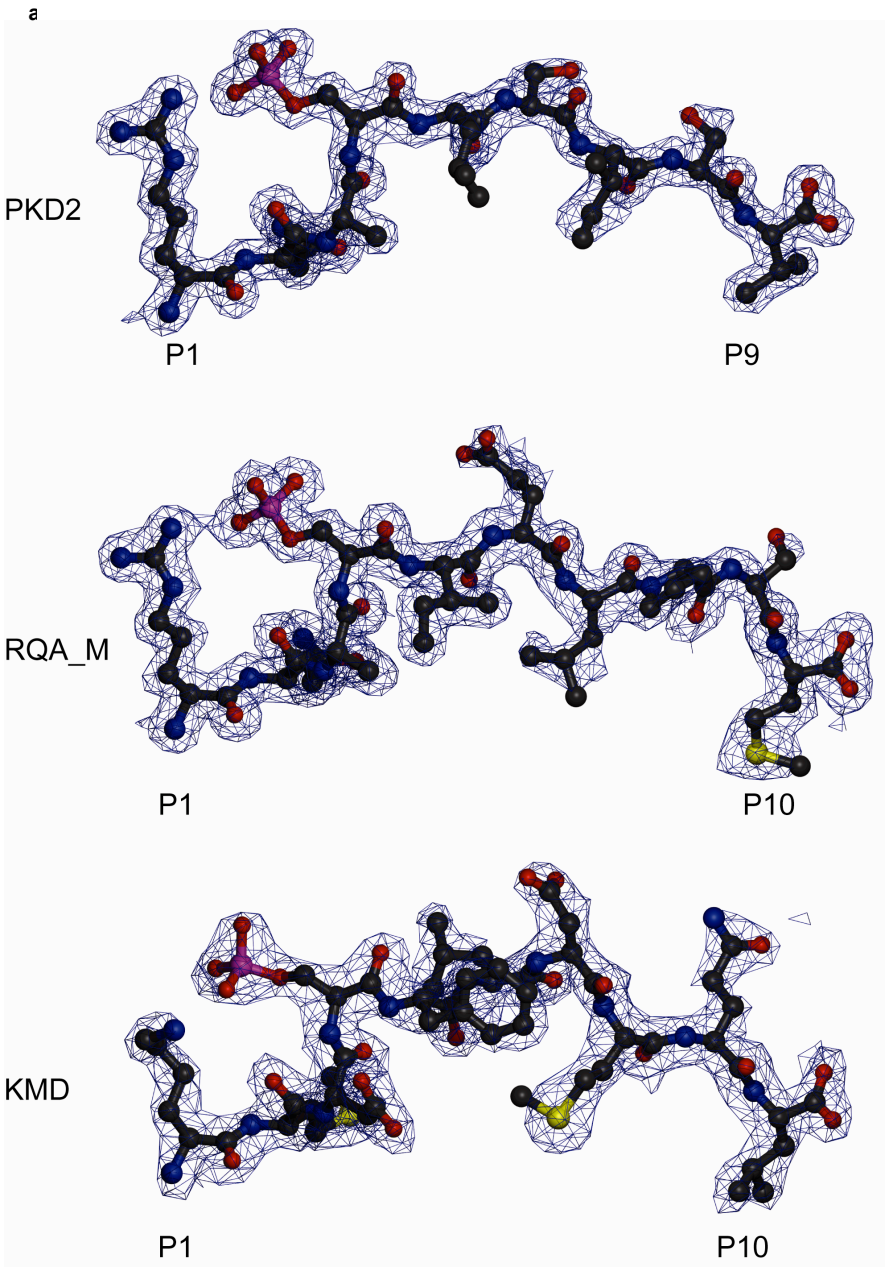
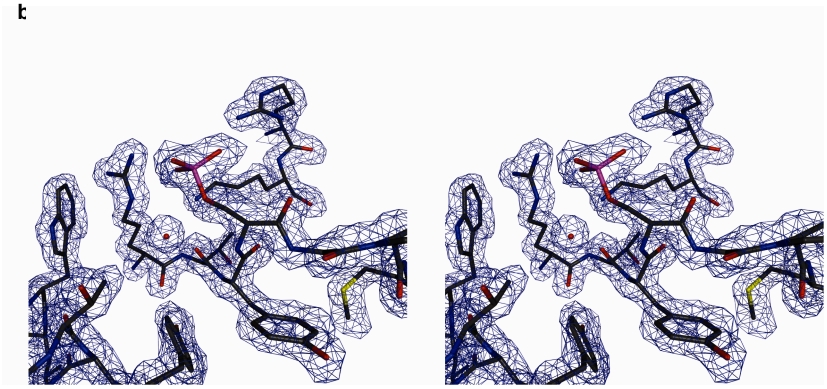

a
PKD2
P1
P9
RQA_M
P1
P10
KMD
P1
P10


b

Supplementary Figure 1. Electron density representation of phosphopeptide structures. (**a**) Structure of PKD2, RQA_M and KMD phosphopeptides, each superimposed on a $2F_o$-F_c electron density map contoured at 1.0 σ (blue wire). Sequences: RQApSLSISV (PKD2), RQApSIELPSM (RQA_M), KMDpSFLDMQL (KMD). (**b**) Stereo diagram of PKD2 model in the region of the P4 phosphorylation site, superimposed on a $2F_o$-F_c electron density map contoured at 1.0 σ, and presented for divergent viewing.

Supplementary Table 1. Sequence anomalies among HLA-A*0201 associated phosphopeptides

Element		Known HLA-A*0201 phosphopeptides[a]	Known non-phosphopeptide HLA-A*0201 ligands		Predicted HLA-A*0201 phosphopeptide ligands	
			Immune Epitope[b]	B-LCL[c]	EMBL[d]	PhosphoSite[e]
Phosphate						
	P4	25 (68%)	-	-	162 (13%)	311 (13%)
	Other	12 (32%)	-	-	1064 (87%)	2058 (87%)
					P<0.0001	*P<0.0001*
P1						
	Arg, Lys	23 (62%)	78 (12%)	46 (8%)	52 (23%)	78 (25%)
	Other	14 (38%)	591 (88%)	505 (92%)	172 (76%)	233 (75%)
			P<0.0001	*P<0.0001*	*P<0.0001*	*P<0.0001*
P2						
	Leu	19 (51%)	383 (57%)	381 (69%)	-	-
	Met, Thr, Val, Gln	18 (49%)	128 (19%)	86 (16%)	-	-
	Other	0 (0%)	158 (24%)	84 (15%)	-	-
			P<0.0001	*P<0.0001*		
PC						
	Val, Leu	26 (70%)	445 (67%)	396 (72%)	-	-
	Other	11 (30%)	224 (33%)	155 (28%)	-	-
			P=0.64	*P=0.8343*		

[a] HLA-A*0201 associated phosphopeptides from Zarling, A.L., et al., *Proc Natl Acad Sci* **103**, 14889 (2006).

[b] peptide dataset of 669 9-10mer HLA-A*0201 binding peptides from www.immuneepitope.org.

[c] peptide dataset of 551 9-10mer naturally processed peptides extracted from HLA-A*0201 molecules associated with the JY B-LCL cell line (unpublished).

[d] dataset of 1,226 nonamer phosphopeptides predicted to bind to HLA-A*0201 based on P2=L and P9=L or V from among 8,770 human p-Ser phosphorylation sites in the EMBL dataset (http://phospho.elm.eu.org/) (Diella, F., et al., *BMC Bioinformatics* **5**, 79 (2004)).

[e] dataset of 2,369 nonamer phosphopeptides predicted to bind to HLA-A*0201 based on P2=L and P9=L or V from among 13,508 human pS phosphorylation sites in the Phosphosite dataset (http://www.phosphosite.org) (Hornbeck, P. V., et al., *Proteomics* **4**, 1551 (2004)).

P values were calculated using a two tailed chi-squared test for the contingency table generated from the A2 phosphopeptide and relevant comparative datasets.

Supplementary Table 2. Ability of peptide binding algorithms to predict HLA-A2 restricted phosphopeptides

	HLA-A2 associated phosphopeptides[a]	HLA-A2 associated non-phosphorylated peptides[b]
BIMAS[c]		
Score ≥ 100	7 (23%)	284 (52%)
Score <100	23 (77%)	267 (48%)
(*P*=0.0026)		
SYFPEITHI[d]		
Score ≥ 20	20 (67%)	460 (83%)
Score <20	10 (33%)	91 (17%)
(*P* =0.0179)		

[a] from Zarling, A.L., et al., *Proc Natl Acad Sci* **103**, 14889 (2006).

[b] dataset of 551 9-10mer naturally processed peptides from B-LCL described in Methods.

[c] HLA-A*0201 9mer and 10mer algorithms and suggested prediction threshold of 100 are from http://www-bimas.cit.nih.gov/.

[d] HLA-A*0201 9mer and 10mer algorithms are from www.syfpeithi.de/. The prediction threshold used was 20.

P values were calculated using a two tailed chi-squared test for the contingency table generated from the phosphopeptide and non-phosphorylated peptide datasets.

Supplementary Table 3. Crystallization, data processing and refinement statistics for phosphopeptide-HLA-A2 complexes

	PKD2	RQA_M	RTY	KMD
Peptide				
Sequence	RQApSLSISV	RQApSIELPSM	RTYpSGPMNKV	KMDpSFLDMQL
Source	Protein kinase D2	Lymphocyte specific protein 1	Premature ovarian failure, 1B	Nedd4 binding protein 2
Crystallization				
Precipitant (%)	PEG 8K (17)	PEG 8K (17)	PEG 3350 (21)	PAA (24)
Buffer (0.1 M) pH 7.5	Hepes	Hepes	-	Hepes
Salt (M)	-	-	NaSCN (0.1)	$MgCl_2$ (0.09)
HLA-A2 (mg/ml)	11	8	15	26
Data Processing				
Resolution (Å)	20-1.6	20-1.65	20-1.7	20-2.2
Unit cell Parameters				
a (Å), b (Å)	117.8, 54.8	112.6, 55.1	117.6, 53.1	57.1, 80.2
c (Å), β (°)	75.8, 105	75.9, 103.5	75.7, 104.8	57.3, 115.9
Space Group	C2	C2	C2	$P2_1$
Total reflections	380435	353305	350718	230491
Unique reflections	59358	50641	46220	22776
Multiplicity	6.4	7.0	7.5	10.1
Completeness (%)	96 (81.7)	92.8 (74.9)	92.6 (64.4)	95.8 (75.6)
R_{merge} (%)	3.3 (44.1)	3.8 (66.4)	4.4 (49)	11.6 (33.3)
$I/\sigma(I)$	29.9 (3.1)	28.5 (2.6)	28.5 (2.7)	16.5 (4.8)
Refinement				
Resolution (Å)	20-1.60	20-1.65	20-1.70	20-2.20
Reflections used	58959	50184	45954	22724
R_{cryst} (%)	20.1	20.0	19.8	22.5
R_{free} (%)	23.2	23.1	22.9	25.6
Protein residues	383	384	385	382
Water molecules	400	368	328	119
RMS deviations				
Bond lengths (Å)	0.012	0.014	0.013	0.013
Bond angles (°)	1.31	1.34	1.42	1.57

Figures in parentheses in the data processing section apply to data in the highest resolution shell.

Supplementary Table 4. Conserved hydrogen bonding interactions between HLA-A2 and the N and C termini of phosphopeptides

Molecular Contacts		PKD2	RQA_M	RTY	KMD
HLA-A2	**Peptide**	**Hydrogen bonding distances (Å)**			
Tyr7 (OH)	P1 (N)	2.9	2.9	3.0	3.0
Tyr171 (OH)	P1 (N)	2.6	2.7	2.8	2.7
Tyr159 (OH)	P1 (O)	2.7	2.7	2.6	2.8
Glu63 ($O^{\varepsilon 1}$)	P2 (N)	3.0	2.9	3.0	3.0
Lys66 (N^{ζ})	P2 (O)	2.9	2.9	2.9	2.9
Tyr99 (OH)	P3 (N)	3.0	3.0	2.9	3.0
Trp147 ($N^{\varepsilon 1}$)	PC-1 (O)	2.8	-	3.0	3.0
Asp77 ($O^{\delta 1}$)	PC (N)	3.0	-	2.9	3.0
Thr143 ($O^{\gamma 1}$)	PC (OXT)	2.7	2.7	2.6	2.9
Tyr84 (OH)	PC (OXT)	2.8	2.7	2.8	3.1
Lys146 (N^{ζ})	PC (OXT)	-	3.0 (PC-1)	3.2	3.1

Supplementary Table 5. Stabilizing interactions involving the p-Ser moiety and HLA-A2

Molecular Contacts		PKD2	RQA_M	RTY	KMD
α1 helix					
HLA-A2	**Peptide**	**Hydrogen bonding distances (Å)**			
Arg65 (NH2)	P4 (O2P)	2.9	2.8	3.0	-
Lys66 (N^{ζ})	P4 (O1P)	2.8	2.8	2.9	2.8
Lys66 (N^{ζ})	P4 (O^{γ})	-	-	-	3.1
Intra-peptide					
P1 Arg (NH2)	P4 (O1P)	2.7	2.9	2.6	-
P1 Arg (N^{ε})	P4 (O1P)	-	-	3.2	-
P1 Lys (N^{ζ})	P4 (O1P)	-	-	-	2.9
H_20 (1)	P2 (O)	2.8	2.9	2.8	-
H_20 (1)	P4 (O^{γ})	3.0	3.0	2.9	-
H_20 (1)	P1 (N^{ε})	2.9	3.0	3.1	-
α2 helix		**Number of hydrophobic contacts**			
Trp167	P1 (Arg)	18	18	22	-
Trp167	P1 (Lys)	-	-	-	15

Supplementary Table 6A. van der Waals and hydrophobic contacts between P2 anchor and HLA-A2

	ILK	KMD	RTY	PKD2	RQA_M
P2 anchor	Leu	Met	Thr	Gln	Gln
	C^{α} (6)	C^{α} (7)	C^{α} (5)	C^{α} (7)	C^{α} (7)
	C^{β} (6)	C^{β} (4)	C^{β} (6)	C^{β} (6)	C^{β} (6)
	$C^{\gamma 1}$ (7)	C^{γ} (9)	$O^{\gamma 1}$ (6)	C^{γ} (8)	C^{γ} (6)
	$C^{\gamma 2}$ (16)	S^{δ} (5)	$C^{\gamma 2}$ (9)	C^{δ} (8)	C^{δ} (6)
	$C^{\delta 1}$ (6)	C^{ε} (14)		$O^{\varepsilon 1}$(8)	$O^{\varepsilon 1}$ (8)
				$N^{\varepsilon 2}$(8)	$N^{\varepsilon 2}$ (8)
Total contacts	41	39	26	45	41

Value in parentheses correspond to the number of hydrophobic and van der Waals contacts mediated between the P2 anchor side chain atom and HLA-A*0201. ILK contacts were determined from the structure of HLA-A2 complexed with HIV-1 reverse transcriptase peptide (PDB code 1HHJ). Two atoms are considered to be in contact if dij<ri+rj+t where dij is the distance between two non-bonded atoms i and j, ri is the van der Waals radius for atom i, rj is the van der Waals radius for atom j and t is the tolerance limit (1Å).

Supplementary Table 6B. Hydrogen bonding interactions mediated between P2 Gln anchor and HLA-A2

Molecular contacts		PKD2	RQA_M
Direct			
Peptide	**HLA-A2**	**Distances (Å)**	
P2 Gln ($N^{\varepsilon 2}$)	Glu63 ($O^{\varepsilon 1}$)	2.9	3.0
P2 Gln ($N^{\varepsilon 2}$)	Glu63 (O)	3.0	3.1
P2 Gln ($O^{\varepsilon 1}$)	H_20 (2)	2.8	2.8
Water mediated			
H_20 (2)	Tyr99 (OH)	2.8	2.9
H_20 (2)	His70 ($N^{\delta 1}$)	2.9	-
H_20 (2)	H_20 (3)	2.7	2.7
H_20 (3)	P3 (O)	2.6	2.6

Supplementary Table 6C. van der Waals and hydrophobic contacts between PC anchor and HLA-A2

	ILK	PKD2	RTY	KMD	RQA_M
PC anchor	Val	Val	Val	Leu	Met
	C^{α} (7)	C^{α} (6)	C^{α} (4)	C^{α} (4)	C^{α} (3)
	C^{β} (5)	C^{β} (5)	C^{β} (6)	C^{β} (6)	C^{β} (5)
	$C^{\gamma 1}$ (13)	$C^{\gamma 1}$ (13)	$C^{\gamma 1}$ (12)	C^{γ} (7)	C^{γ} (7)
	$C^{\gamma 2}$ (10)	$C^{\gamma 2}$ (10)	$C^{\gamma 2}$ (12)	$C^{\delta 1}$ (11)	S^{δ} (6)
				$C^{\delta 2}$ (12)	C^{ε} (7)
Total contacts	35	34	34	40	28

Value in parentheses correspond to the number of hydrophobic and van der Waals contacts mediated between the PC anchor side chain atom and HLA-A*0201. ILK contacts were determined from the structure of HLA-A2 complexed with HIV-1 reverse transcriptase peptide (PDB code 1HHJ). Van der Waals contacts calculated as in Supplementary Table 6a.