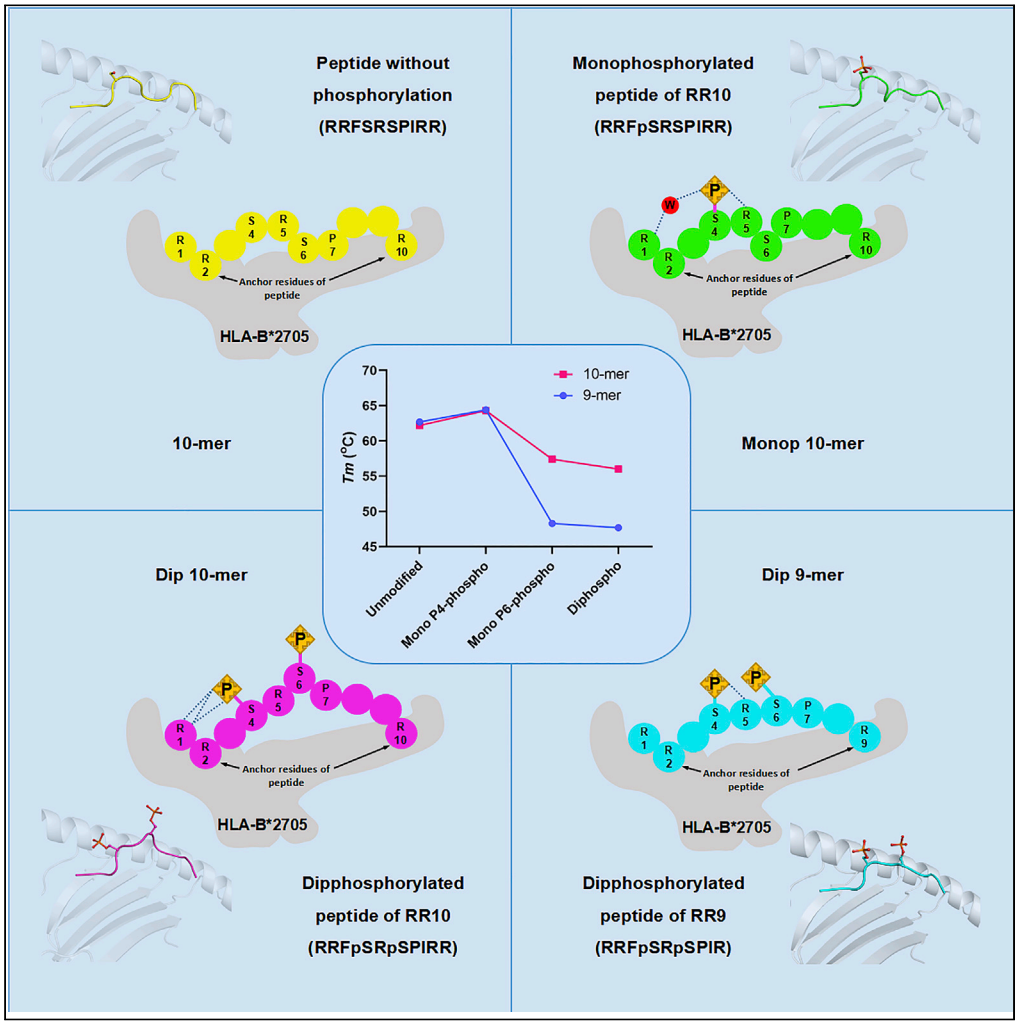
iScience

CellPress
OPEN ACCESS

Article

Phosphosite-dependent presentation of dual phosphorylated peptides by MHC class I molecules



Yingze Zhao, Mingwei Sun, Nan Zhang, ..., Catherine C.L. Wong, George F. Gao, William J. Liu

gaof@im.ac.cn (G.F.G.)
liujun@ivdc.chinacdc.cn (W.J.L.)
catherine_wong@bjmu.edu.cn (C.C.L.W.)

Highlights

Diphosphorylation with an interval is prevalent among multiphosphorylated peptides

Diphosphopeptide presentation is HLA specific and different from monophosphopeptide

The cooperativity of peptide conformation within P4 and P6 phosphorylation sites

The diphosphorylation attenuates peptide binding affinity to HLA-B27

Zhao et al., iScience 25, 104013
April 15, 2022
https://doi.org/10.1016/j.isci.2022.104013

iScience

CellPress
OPEN ACCESS

Article

Phosphosite-dependent presentation of dual phosphorylated peptides by MHC class I molecules

Yingze Zhao,[1,2,7] Mingwei Sun,[3,7] Nan Zhang,[4,7] Xueyuan Liu,[1,5] Can Yue,[1,6] Lei Feng,[1] Shushen Ji,[4] Xiao Liu,[4] Jianxun Qi,[3] Catherine C.L. Wong,[4,*] George F. Gao,[1,2,3,6,*] and William J. Liu[1,2,8,*]

SUMMARY

Phosphopeptides presented by major histocompatibility complex (MHC) class I have been regarded as a pivotal type of cancer neoantigens that are recognized by T cells. The structural basis of single-phosphorylated peptide presentation has been well studied. Diphosphorylation with one interval between two sites is one of the prevalent forms of multisite-phosphorylated peptides. Herein, we determined the molecular basis of presentation of two P4/P6 double pS-containing peptides by HLA-B27 and compared them with unmodified and single-phosphorylated peptide complexes. These data clarified not only the HLA allele-specific presentation of phosphopeptides by MHC class I molecules but also the cooperativity of peptide conformation within P4 and P6 phosphorylation sites. The phosphorylation of P6 site can influence the binding mode of P4 phosphorylated site to HLA-B27. And we found the diphospho-dependent attenuated effect of peptide binding affinity. This study provides insights into the MHC presentation features of diphosphopeptides, which is different from monophosphopeptides.

INTRODUCTION

The cytotoxic response and cytokine release of $CD8^+$ T cells facilitated by T cell receptor (TCR)-mediated specific recognition of major histocompatibility complex-binding peptides (MHC-binding peptides) are important for antiviral and anti-tumor immunity. During viral infection or carcinogenesis of cells, alterations in protein expression and metabolism can result in modifications of peptide libraries and subsequently lead to the presentation of "altered self" antigens by MHC molecules (Engelhard et al., 2006). It has been demonstrated that posttranslationally modified peptides, including phosphopeptides, could diversify the immunopeptidome and have thus become potential targets for T cell recognition (Doyle and Mamula, 2001; Mohammed et al., 2008). Phosphorylation caused by abnormal protein kinase activity has gained increasing attention in the treatment of tumors because it often serves as a marker for malignant transformation, activation of oncogenic signaling pathways, cell growth, and differentiation (Cohen, 2000). The mutation-induced abnormal regulation of a limited number of kinases often leads to the activation of some signaling pathways and increases the level of protein phosphorylation in cells (Huang et al., 2018). Multisite phosphorylation may also collectively lead to transient changes in catalytic activity, structural properties, protein circulation and interaction, and cross-linked clustering of lipids and cell migration. Thus, changes in protein phosphorylation in malignant tumors manifest as a distinctive "altered self," which also serves as immune signals (Eggleton et al., 2008).

Previous studies have found that approximately 30% of cellular proteins (from yeasts to humans) may contain phosphorylated Ser, Thr, or Tyr residues (Schweiger and Linial, 2010). Most phosphopeptides are derived from phosphorylated proteins that have been transported into the ER after the processing by the proteasome along with other components of the peptide loading complex (Mohammed et al., 2008). T cells can recognize peptides based on their phosphate groups and can distinguish phosphopeptides from nonphosphopeptides (Mohammed et al., 2017). For instance, HLA-A2-restricted, tumor-specific phosphopeptides exhibit immunogenicity, which can be recognized by cytotoxic T lymphocytes (CTLs) in HLA-A2 transgenic mice (Zarling et al., 2006). A series of phosphopeptides originating from proteins associated with cell-cycle regulation and oncogenic signaling pathways have been identified on MHC I molecules in different tumor cell lines. For instance, Zarling et al. identified 36 HLA-A2-restricted

[1]NHC Key Laboratory of Biosafety, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention (China CDC), Beijing 100052, China

[2]Research Unit of Adaptive Evolution and Control of Emerging Viruses (2018RU009), Chinese Academy of Medical Sciences, Beijing 102206, China

[3]CAS Key Laboratory of Pathogen Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences (CAS), Beijing 100101, China

[4]Center for Precision Medicine Multi-Omics Research, Peking University Health Science Center, Peking University, Beijing 100191, China

[5]School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, China

[6]Savaid Medical School, University of Chinese Academy of Sciences, Beijing 100049, China

[7]These authors contributed equally

[8]Lead contact

*Correspondence: gaof@im.ac.cn (G.F.G.), liujun@ivdc.chinacdc.cn (W.J.L.), catherine_wong@bjmu.edu.cn (C.C.L.W.)

https://doi.org/10.1016/j.isci.2022.104013

phosphopeptides on melanoma cells, ovarian cancer cells, and B lymphoblastoid cells in a comparative study (Zarling et al., 2006). The structures of seven of these phosphopeptides complexed to HLA-A2 were subsequently resolved by Mohammed et al. and Petersen et al., which revealed the special presentation properties of HLA-A2-restricted phosphopeptides (Mohammed et al., 2008, 2017; Petersen et al., 2009). The structures of HLA-A2-P4/P5 phosphopeptide complexes demonstrated the molecular mechanism underlying the presentation of tumor-specific "altered self" epitopes.

The increasing amount of available mass spectrometry data and the improvement and advancement of phosphorylation assays provide more opportunities for studying the characteristics of phosphosites. Multisite phosphorylation is a common phenomenon, and the distribution of phosphosites is inversely proportional to the distance between two phosphosites, among which proximal phosphorylation with a distance of two amino acid residues is the most common phenomenon (Schweiger and Linial, 2010). In 2000, Zarling et al. identified (via mass spectrometry) HLA-A2-, A3-, B7-, and B27-restricted phosphopeptides that can be naturally processed and presented on numerous cell lines (Zarling et al., 2000). B27-restricted phosphopeptide (RRFpSRpSPIRR, derived from RNA-binding proteins, i.e., mRNA splicing cofactor, Son3) has two phosphorylated Ser residues at the P4 and P6 positions. Moreover, a natural 9-residue truncated analog (RRFpSRpSPIR) and a monophosphorylated form (RRFpSRSPIRR) of the phosphopeptide have been identified within the same elution fraction of the liquid phase. However, although this is a more common form of phosphoprotein degradation products, the molecular basis for the presentation of diphosphopeptides is still unclear.

In this study, we resolved the structures of HLA-B*2705 complexes with these two (10- and 9-residue) diphosphopeptides, together with the monophosphopeptide and peptide without any phosphorylation via X-ray crystallography. Our results showed the unique presentation properties of diphosphopeptides compared with monophosphopeptides. In addition, by comparing the diphosphopeptide, P4 monophosphopeptide, with nonmodified peptide structures, we further clarified the molecular basis for the presentation of self-antigens as neoepitopes with modifications. The thermal stabilities of pMHC complexes formed with diphosphorylated, monophosphorylated, and nonmodified peptides were determined via circular dichroism (CD) spectroscopy, and we showed that phosphate groups exhibit phosphosite-dependent effects on the peptide-binding affinity and the stability of pMHC. These results also provide a structural and thermodynamic basis for the presentation of diphosphorylated peptides as an additional type of posttranslationally modified antigens.

RESULTS

Diphosphorylation at an interval of one amino acid residue as one of the main forms of protein phosphorylation

To analyze diphosphorylation patterns on tumor-associated proteins, we acquired the location information of human protein phosphorylation sites and related cancer information through PhoSigNet database (http://119.3.70.71/PhoSigNet/) (Zhang et al., 2015). The adjacent diphosphorylation were defined as PP and the diphosphorylation with one or two interval residues as P1P or P2P. Our analysis of tumor-associated proteins in different diseases showed that the average number of diphosphorylation with one interval residue among tumor-associated proteins (average P1P number/protein) is much higher than the proteins that are currently known not associated with cancer (Figure 1A). Generally, the proteins from the digestive-tract-associated cancers such as esophageal, gastric, colorectal cancers possess a higher level of phosphorylation. Statistical analyses showed that the distribution of phosphosites is inversely proportional to the distance between two phosphosites, among which proximal phosphorylation between two residues is the common phenomenon (Figure 1B). Furthermore, previous studies showed that the phosphopeptides contained a positively charged amino acid (Arg or Lys) at position 1, and p-Ser or p-Thr residue at position 4 was relatively common for monophosphopeptide, which can be recognized as the RXXpSXXXXX motif (for 9-mer peptide) (Mohammed et al., 2008; Zarling et al., 2006). So, according to the location information of phosphorylation site of each protein in the PhoSigNet database, we analyzed the frequency of the third amino acid before the P1P phosphorylation form, and the result showed that the positively charged amino acid Arg had the highest frequency (Figure 1C). This result demonstrated a common motif RXXpSXpSXXX for 9-mer biphosphorylated peptides.

Overall structures of diphosphopeptide–HLA-B*2705 complexes

In this study, we resolved the crystal structures of HLA-B*2705 complexes with 10-residue diphosphopeptide Dip-RR10 (RRFpSRpSPIRR, where "pS" indicates phosphorylated P4- and P6-Ser residues), 9-residue

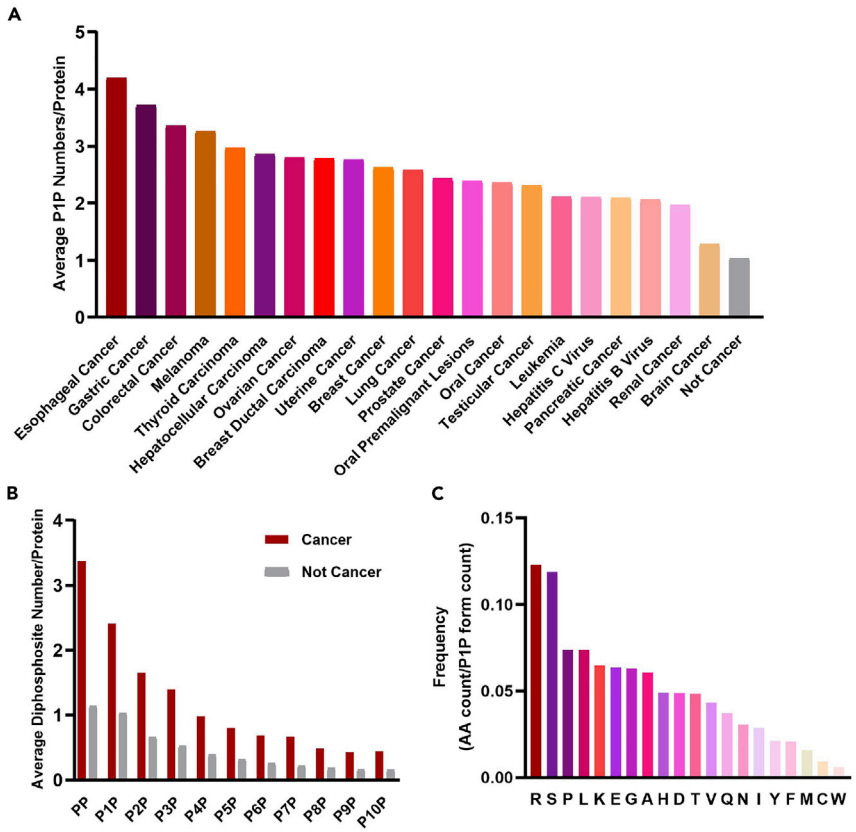


Figure 1. Distribution of diphosphosites

According to the diphosphorylation sites information of human protein and related cancer information from PhoSigNet database.

(A) The average number of diphosphorylation with one interval residue among tumor-associated proteins (average P1P number/protein) in different diseases.

(B) The average number of different diphosphorylation patterns on tumor-associated proteins or non-tumor-associated proteins is counted. The adjacent diphosphorylation was defined as PP, and the diphosphorylation forms with spaced amino acids were defined as PXP, etc. (X = 1, 2, 3, …).

(C) The frequency of the third different amino acid in front of the P1P phosphorylation form.

diphosphopeptide Dip-RR9 (RRFpSRpSPIR), 10-residue P4 monophosphopeptide Mop-RR10 (RRFpSRSPIRR), and nonmodified 10-residue peptide RR10 (RRFSRSPIRR) (Table 1). The structures of these HLA-B*2705 complexes retained the overall structural characteristics of typical pMHC complexes (Figures 2A and 2D). All peptides with well-defined electron densities (Figure S1) displayed extended conformations in the peptide-binding grooves. Both P2 and Pc positions in these four peptides are positively charged Arg residues that serve as primary anchor residues to embed side chains into pocket B and pocket F of HLA-B*2705 molecules (Figures 2B and 2E), respectively. Peptides in these four HLA-B*2705 complexes resolved in this study have structurally identical directions of side-chain extension at the P1–P4 residues (Figure 2G). In contrast, these peptides displayed different directions of side-chain extension at the P5–P9 residues with a higher degree of flexibility and variability (Figure 2G). Tqhe phosphate groups attached to the hydroxyl group of monophosphopeptide Monop-RR10, diphosphopeptide Dip-RR10 and Dip-RR9 pointing upward allow direct contact between their dominant antigenic sites with corresponding TCRs. This is similar to the previously resolved HLA-A2-restricted phosphopeptides, which are also phosphorylated at the side chain of the P4-Ser residue. Moreover, the P6 phosphate group on diphosphopeptides displayed even more pronounced solvent exposure and became the apex of the phosphopeptide backbone (Figures 2C and 2F).

Site-specific phosphorylation affects peptide conformation

To compare the conformational differences of the nonmodified peptide (RR10) and its corresponding P4 phosphopeptide (Monop-RR10), we determined both the structures of HLA-B*2705 complexed to the

Table 1. Data processing and refinement statistics

Parameter	HLA-B*2705/RR10	HLA-B*2705/Monop-RR10	HLA-B*2705/Dip-RR10	HLA-B*2705/Dip-RR9
Data processing				
Space group	$P2_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell parameters				
a (Å)	43.40	50.73	50.83	51.20
b (Å)	52.35	82.24	82.30	82.40
c (Å)	93.21	108.44	109.45	110.00
α (°)	90.00	90.00	90.00	90.00
β (°)	93.23	90.00	90.00	90.00
γ (°)	90.00	90.00	90.00	90.00
Wavelength (Å)	1.00000	1.00000	1.54178	1.54178
Resolution (Å)	50.00–1.60 (1.66–1.60)	50.00–1.80 (1.86–1.80)	50.0–2.0 (2.07–2.00)	50.00–2.10 (2.18–2.10)
Total reflections	203,885	303,959	215,220	266,575
Unique reflections	55,528	41,560	31,363	27,813
Redundancy	3.7 (2.6)	7.3 (7.4)	6.8 (6.5)	9.6 (8.1)
Completeness (%)[a]	99.5 (96.6)	98.5 (96.9)	97.8 (81.9)	99.4 (95.3)
R_{merge} (%)[b]	7.9 (41.3)	11.1 (51.7)	8.9 (36.6)	11.2 (57.7)
I/σ	16.03 (1.90)	17.67 (3.82)	20.89 (4.84)	20.00 (3.11)
Refinement statistics				
R_{work} (%)[c]	17.17 (22.44)	17.65 (22.03)	16.84 (18.71)	18.15 (22.75)
R_{free} (%)	20.93 (24.91)	20.78 (27.38)	20.78 (24.10)	22.16 (29.04)
RMSD				
Bonds (Å)	0.007	0.007	0.007	0.008
Angles (°)	1.19	1.17	1.19	1.18
Average B factor (Å)	20.53	24.75	25.25	33.81
Ramachandran plot quality				
Favored (%)	97.37	97.61	98.13	98.66
Allowed (%)	2.37	2.39	1.87	1.34
Outliers (%)	0.26	0	0	0

[a]Data completeness = (number of independent reflections)/(total theoretical number).
[b]$R_{merge} = \sum_{hkl}\sum_i|I_i - \langle I\rangle|\sum_{hkl}\sum_i I_i$, where I_i is the observed intensity and $<I>$ is the average intensity of multiple observations of symmetry-related reflections.
[c]$R = \sum_{hkl}||F_{obs}| - k|F_{cal}||/\sum_{hkl}|F_{obs}|$, where R_{free} is calculated for a randomly chosen 5% of reflections, and R_{work} is calculated for the remaining 95% of reflections used for structure reflnement.

two peptides, although RR10 was not found in the naturally eluded products from the cell lines. The extended conformation of peptide RR10 is tightly bound to the peptide-binding groove and gives rise to the M-shaped conformation of the peptide backbone with a P6-Ser as a secondary anchoring residue in pocket C (Figures 3A and 3G). Compared with peptide RR10, the phosphorylated P4-Ser residue of peptide Monop-RR10 does not significantly alter the conformation at the P4 position or its adjacent residues. However, there is an unexpected rigid torsion at the P7–P8 positions (Figures 3A and 3B). There were approximately 110 and 60° rotations at the C–Cα bonds of P6-Ser and P8-Ile residues, respectively. The side chain of the P6-Ser residue of Monop-RR10 is more deeply embedded into the peptide-binding grooves, leading to an even more pronounced bulged conformation at the P4 and P5 positions. In addition to the phosphorylated P4-Ser residue in the Monop-RR10 peptide, there is a second phosphorylation at the P6 position in the Dip-RR10 peptide, and the conformational effects of this modification are mostly observed at the middle portion of the peptide (P5–P8 positions) during antigen presentation. In the peptide-binding grooves, P4 phosphorylation gives rise to peptides with an inward conformation. This is different from P6-Ser phosphorylation, which results in a significantly bulged conformation at the P5–P7 residues (including the phosphosite), leading to the greatest amount of solvent exposure and potential immunogenicity (Figures 3C and 3D). Dip-RR9 was considered a product of a C-terminal truncated form

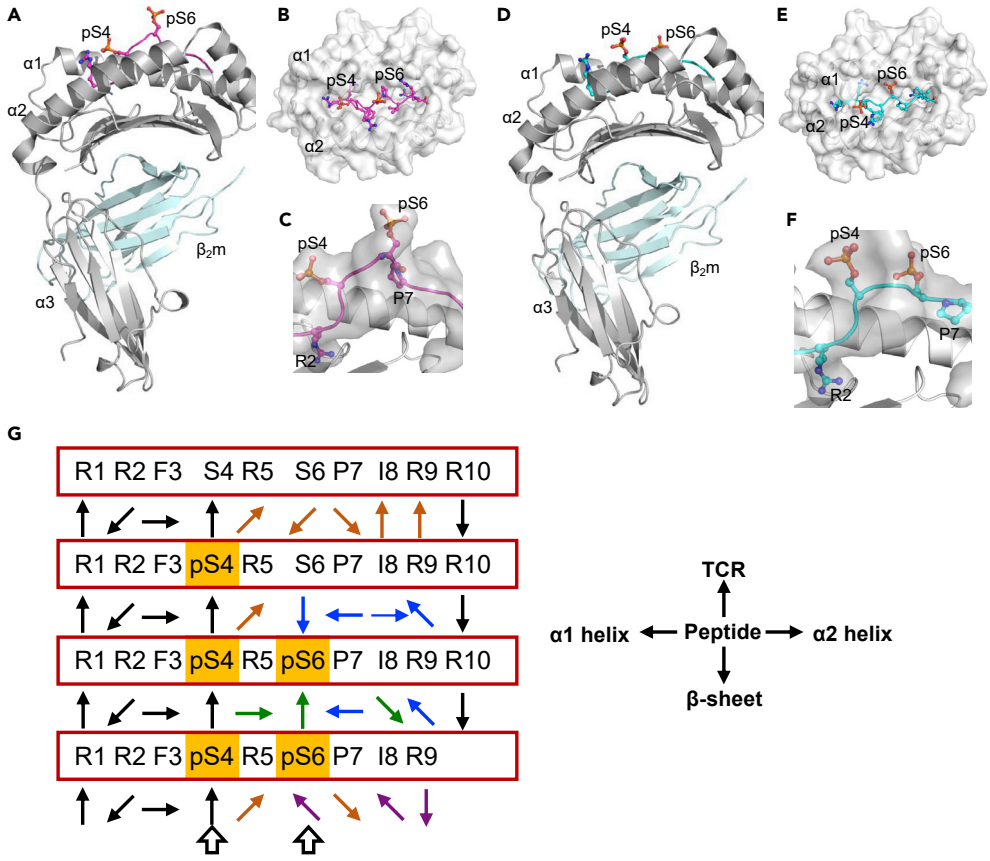


Figure 2. Overall structures of the HLA-B*2705 complex with diphosphopeptides

(A and D) HLA-B*2705 binds to the diphosphopeptides, Dip-RR10 (A) and Dip-RR9 (D).

(B and E) Top view of diphosphopeptides Dip-RR10 (B) and Dip-RR9 (E) in stick and sphere models in purple and cyan colors, respectively. The HLA-B*2705 is shown with white surfaces.

(C and F) Dip-RR10 (C) and Dip-RR9 (F) extend the P4 and P6 phosphosites (pS4 and pS6) outside the binding grooves for solvent exposure. Peptides are shown as gray surfaces. The red spheres represent oxygen atoms, the blue spheres represent nitrogen atoms and the orange spheres represent phosphorus atoms.

(G) Extension direction of four peptide side chains relative to HLA-B*2705 and TCR as observed from the N terminal to the C terminal (refer to the directional markers on the top of the figure). The arrowheads show the direction of the peptide. See also Figure S1.

of Dip-RR10. Deletion of the C-terminal Arg residue gives rise to a conformation at the P9-Arg residue that is almost identical to that of the P10-Arg residue in the 10-residue peptide and serves as the Pc anchor residue (Figures 3E and 3F). Dip-RR9, which is one residue shorter than Dip-RR10, has a slightly flattened conformation with conformational repositioning of the P4–P8 residues in the peptide-binding grooves. Taken together, this suggests that peptides with the same amino acid sequences may be presented on MHC I molecules with different structural properties and antigenicity following modifications, such as P4- and P6-phosphorylation and C-terminal truncation (Figure S2).

Phosphorylation not only triggers conformational changes in the peptide backbone but also affects the binding mode of amino acid residues in the middle portion of peptides to the peptide-binding grooves of MHC I molecules. The typical antigen-binding mode of MHC I molecules is retained in the structure of HLA-B*2705/RR10 with no peptide modifications. P3-Phe, P6-Ser, and P7-Pro, which are secondary anchor antigen residues, occupy pockets D, C, and E of peptide-binding grooves, respectively (Figure 3G). However, the addition of a phosphate group to the P4-Ser residue in Monop-RR10 does not lead to a significant conformational change in proximal amino acid residues but causes a dramatic rearrangement of amino acid residues in the middle portion of the peptide. The most significant aspect of the rearrangement of amino acid residues is that the side chain of the P6-Ser residue binds closer toward the bottom of the

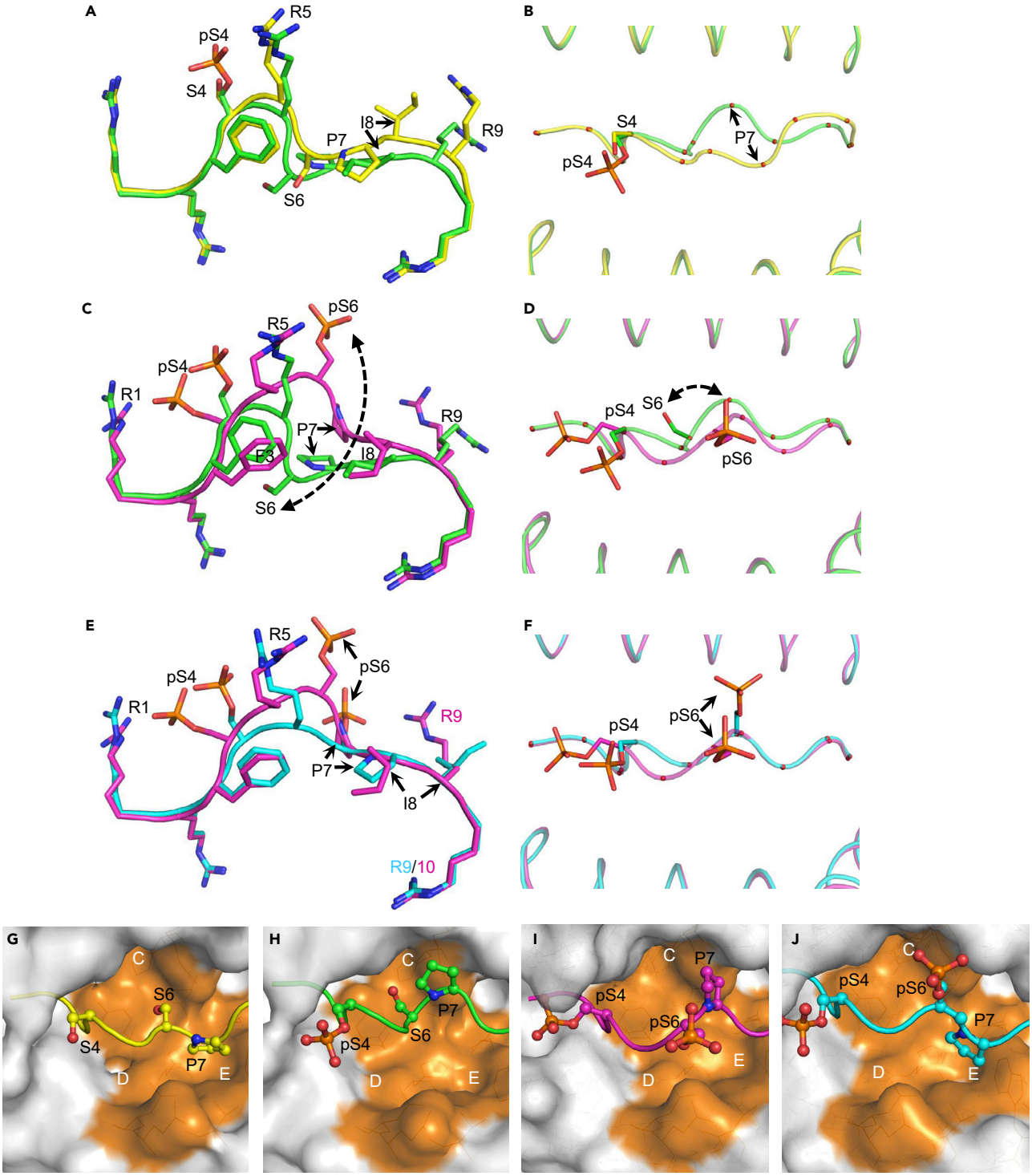


Figure 3. Conformational impacts of diphosphopeptides

(A–F) Pairwise overlapping comparison of conformations of the four peptides of RR10 (yellow), Monop-RR10 (green), Dip-RR10 (purple), and Dip-RR9 (cyan) at the back of the α1-helices in the HLA-B*2705 structures: (A, B) peptide RR10 without phosphorylation and monophosphorylated peptide Monop-RR10; (C, D) monophosphorylated peptide Monop-RR10 and diphosphorylated peptide Dip-RR10; and (E, F) diphosphorylated 10-mer Dip-RR10 and diphosphorylated 9-mer Dip-RR9. Panels B, D, and F correspond to the top view of the orthogonal models of A, C, and E, respectively.

(G–J) Conformational shifts of peptides RR10 (G), Monop-RR10 (H), Dip-RR10 (I), and Dip-RR9 (J) at the peptide-binding grooves of C, D, and E shown with orange surfaces. The side chains of residues P4-Ser, P6-Ser, and P7-Pro and the phosphate groups are shown as sphere-and-stick models. See also Figure S2.

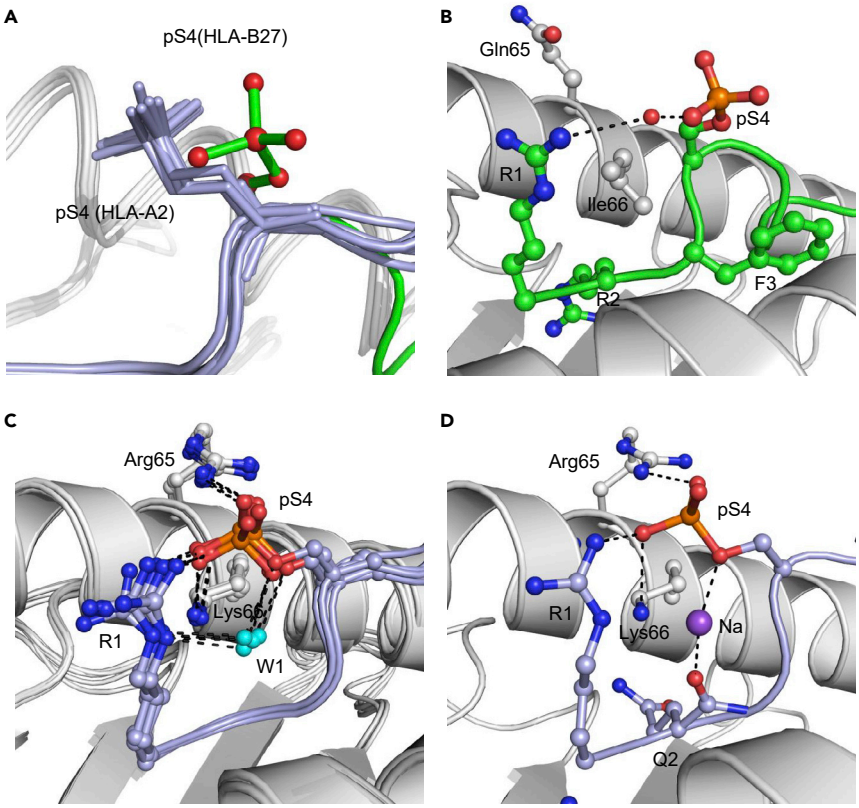


Figure 4. HLA-type-specific interaction of monophosphorylated P4-Ser with surrounding residues

(A) Comparison of the HLA-A*0201-presented phosphopeptides (PDB codes: 3BGM, 2BH8, 3BH9, 4NNX, 4NO2, and 4NO3) and HLA-B*2705/Monop-RR10, showing the distinct conformation of the monophosphopeptide Monop-RR10 presented by HLA-B*2705.

(B) Interactions between monophosphorylated P4-Ser in HLA-B*2705/Monop-RR10 through a water molecule in red and surrounding amino acids on the α1-helix.

(C) The P4-pSer forms a hydrogen bond network between Arg65 and Lys66 of HLA-A*0201 molecules (PDB codes: 3BGM, 2BH8, 3BH9, 4NNX, and 4NO3) and water molecules.

(D) Individual HLA-A*0201 molecules interact with P4-pSer through hydrogen bonds with sodium ions (PDB code: 4NO2). Side chains of the residues are shown as sphere-and-stick models, the red spheres represent oxygen atoms, the blue spheres represent nitrogen atoms, cyan spheres represent water molecules, violet spheres represent sodium atoms, and black-dotted lines represent hydrogen-bonding interactions. See also Figure S3.

peptide-binding grooves. In addition, compared with pocket E of HLA-B*2705/RR10, which is occupied by P7-Pro of the peptide, the side chain rotation of the P7-Pro residue along the axis of the peptide Monop-RR10 backbone leads to the relocation of the pyrrole ring of the P7-Pro residue into pocket C of HLA-B*2705 (Figure 3H). Owing to the steric hindrance of the phosphorylated P6-Ser residue in the diphosphorylated Dip-RR10 structure, the side chain of the phosphorylated Ser residue flips outward the peptide binding groove (Figure 3I). Thus, rotation of the peptide backbone at the P6 position causes the P6 residue to lose its role as an anchor residue in pocket C (Figure 3I). In the HLA-B*2705/Dip-RR9 structure, deletion of one C-terminal amino acid residue in the diphosphorylated 9-residue Dip-RR9 results in the occupation of pocket E by the P7-Pro residue due to the extension of the C-terminus of Dip-RR9 (Figure 3J).

HLA specificity of monophosphopeptides interacting with peptide-binding grooves

Previous structural biology studies resolved the structures of 10 HLA-A*02-restricted phosphopeptides (Mohammed et al., 2008, 2017; Petersen et al., 2009), eight of which had a P1-Arg or Lys motif and a phosphorylated P4-Ser residue. However, the Monop-RR10 peptide presented on HLA-B*2705 molecules does not display structural properties that were observed in the previously resolved phosphoantigen–HLA-A*0201 complexes (Figure 4A). The ionic bonding between the positively charged P1 amino acid residue

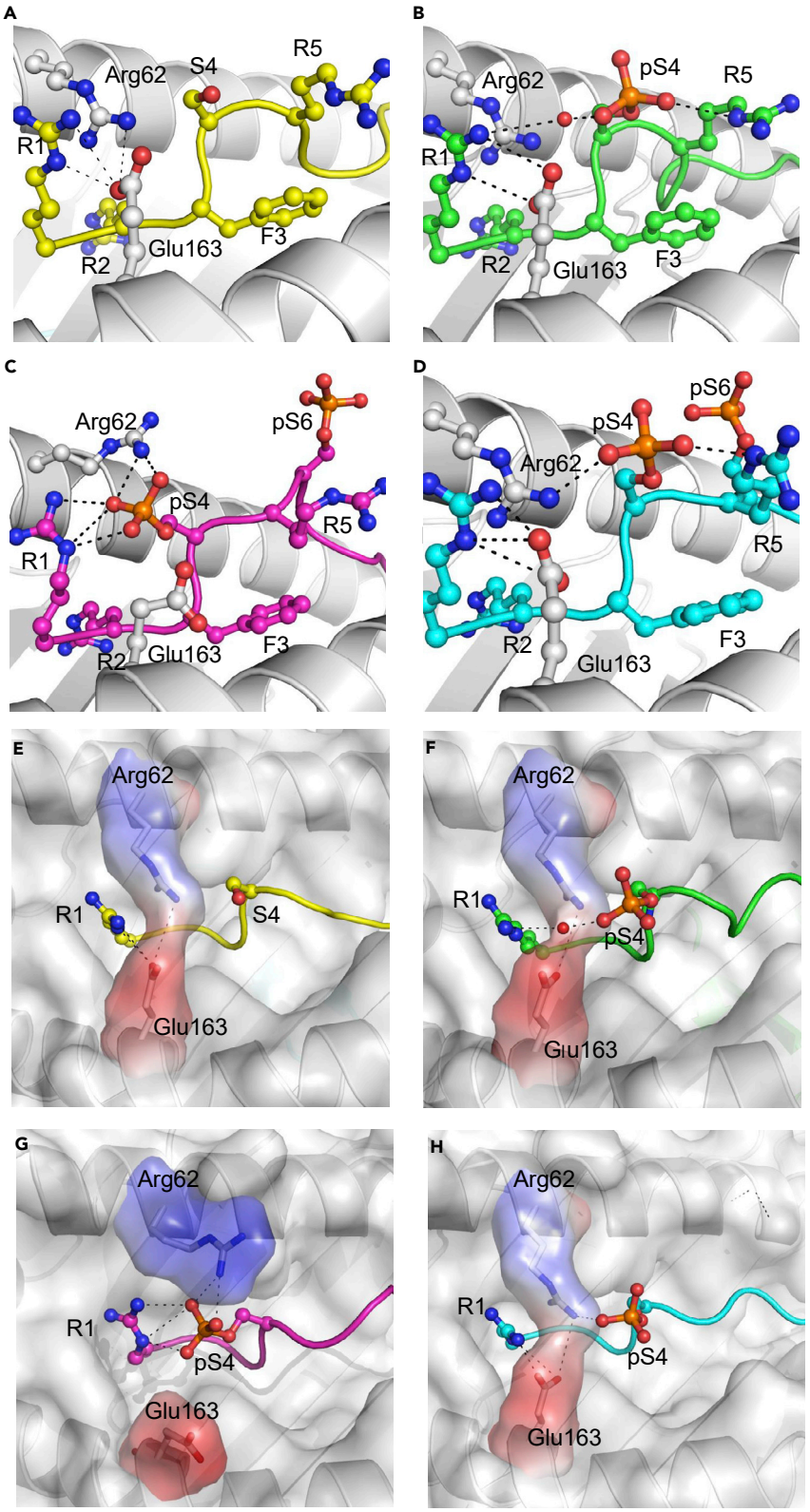

A
Arg62
S4
R5
R1
F3
Glu163
R2
B
Arg62
pS4
R5
R1
R2
Glu163
F3
C
Arg62
pS6
pS4
R5
R1
R2
Glu163
F3
D
pS6
pS4
Arg62
R5
R2
F3
Glu163
E
Arg62
R1
S4
Glu163
F
Arg62
R1
pS4
Glu163
G
Arg62
R1
pS4
Glu163
H
Arg62
R1
pS4
Glu163

Figure 5. Diphosphorylation site-dependent interaction between P4-pSer and its surrounding residues

(A–D) Structure of HLA-B*2705 complexed to the four peptides: (A) RR10 peptide without phosphorylation.

(B) The monophosphopeptide, Monop-RR10.

(C) The diphosphorylated 10-mer peptide, Dip-RR10.

(D) The diphosphorylated 9-mer peptide, Dip-RR9. The peptide side chains are shown as sphere-and-stick models, the red spheres represent oxygen atoms, the blue spheres represent nitrogen atoms, orange spheres represent phosphorus atoms, and black-dotted lines represent hydrogen-bonding interactions.

(E–H) Effects of different phosphorylations on the salt bridges formed between Arg62 on the α1 and Glu163 on the α2 helix in the four complexes: HLA-B*2705/RR10 (E), HLA-B*2705/Monop-RR10 (F), HLA-B*2705/Dip-RR10 (G), and HLA-B*2705/Dip-RR9 (H). In HLA-B*2705/Dip-RR10 (G), the direct interactions between phosphosite P4-Ser and Arg62 open the salt bridge.

and P4-pSer residue of the phosphoantigens presented by HLA-A*0201 was not observed in the structure of the HLA-B*2705/Monop-RR10 complex (Figure 4B). The bonding of these two amino acid residues within the peptide was replaced by water-mediated hydrogen bonding interactions. Furthermore, the phosphate group was also linked to the Nε atom of the P5-Arg residue via hydrogen bonds (Figures 4B and 5). These interactions directly and indirectly maintain the spatial stability of the phosphate group. The positively charged Arg65 and Lys66 residues in the α1-helix of the HLA-A*0201 heavy chain form multiple salt bridges with negatively charged phosphate groups, thus stabilizing its binding to phosphopeptides (Figures 4C and 4D) (Mohammed et al., 2008; Petersen et al., 2009). In contrast, the corresponding residues in HLA-B*2705, Gln65 and Ile66, have no direct interaction with P4-pSer in Monop-RR10. Previously reported interactions between phosphosites (which serve as "surface anchor residues") and peptide-binding grooves were not observed between HLA-B*2705 and P4 monophosphopeptides. Thus, the presence of allele-specific amino acid residues in the peptide-binding grooves of HLA-B*2705 and HLA-A*0201 specifically affect the binding mode of P4 monophosphopeptides to MHC I (Figures S3).

Diphosphosite-dependent interactions between peptides and MHC I molecules

In the HLA-B*2705/RR10 structure, there is no direct interaction between the nonphosphorylated P4-Ser residue and Arg62 residue in the HLA-B*2705/RR10 complex, whereas Arg62 and Glu163 residues form a conventional salt bridge in HLA-B*2705 (Figures 5A and 5E). Meanwhile, the Arg62 residue forms two hydrogen bonds with the oxygen atom in the hydroxyl group of the P2-Arg residue on the peptide main chain. These interactions are quite similar in the HLA-B*2705/Monop-RR10 complex, whereas no direct interaction of Arg62 with the phosphate group was found (Figures 5B and 5F).

In the HLA-B*2705/Dip-RR10 structure, phosphorylation of the P6-Ser residue leads to rotation of the peptide backbone and the flipping up of its side chain and phosphate group, resulting in a prominent bulged conformation at the P5–P7 positions in the peptide backbone. These conformational changes indirectly result in the projection of P4-pSer toward the P1 residue. O2P and O3P atoms in the phosphate group form a stable hydrogen-bonding network with Nη2 and Nε atoms in the P1-Arg residue, respectively. In addition, the Arg62 residue repositions its side chain (the Nη1 atom) and forms two ionic bonds with the phosphate group (O1P and O2P atoms) (Figures 5C and 5G). In contrast, the salt bridge between Arg62 and Glu163 residues of HLA-B*2705 is broken down (Figures 5C and 5G). Moreover, the side chains of the P1-Arg and Glu163 residues, which initially bind to one another via hydrogen bonds in the HLA-B*2705/Monop-RR10 complex, point in opposite directions in the Dip-RR10 complex. In addition, repositioning of the P4-pSer N-terminus prevents the formation of hydrogen bonds with the P5-Arg residue in the diphosphoantigen.

The diphosphopeptide, Dip-RR9, is a one amino acid-shorter peptide at the C-terminus than Dip-RR10 but still has a P9-Arg anchor, and it is presented as a natural antigen on HLA-B27. Dip-RR9 has a relatively flattened conformation in the middle portion (P4–P7 positions) of the peptide (Figures 5D and 5H). The stretch from the C terminus of the peptide brings the P4-pSer residue of Dip-RR9 to a position as in the HLA-B*2705/Monop-RR10 complex. Both Dip-RR9 and Monop-RR10 complexes have almost identical relative amino acid positions around the P4 phosphate group. However, unlike monophosphorylated Monop-RR10, an ionic bond is formed between the O2P atoms of Arg62 and the phosphate group of Dip-RR9 (Figures 5D and 5H).

Structural properties of P6 phosphosites in peptides and the solvent exposure

The bulged conformation of the peptide backbone of the Dip-RR10 complex lifts the P6-pSer residue to become the apex of the antigen presentation plane of the pMHC molecule. There is no observed MHC

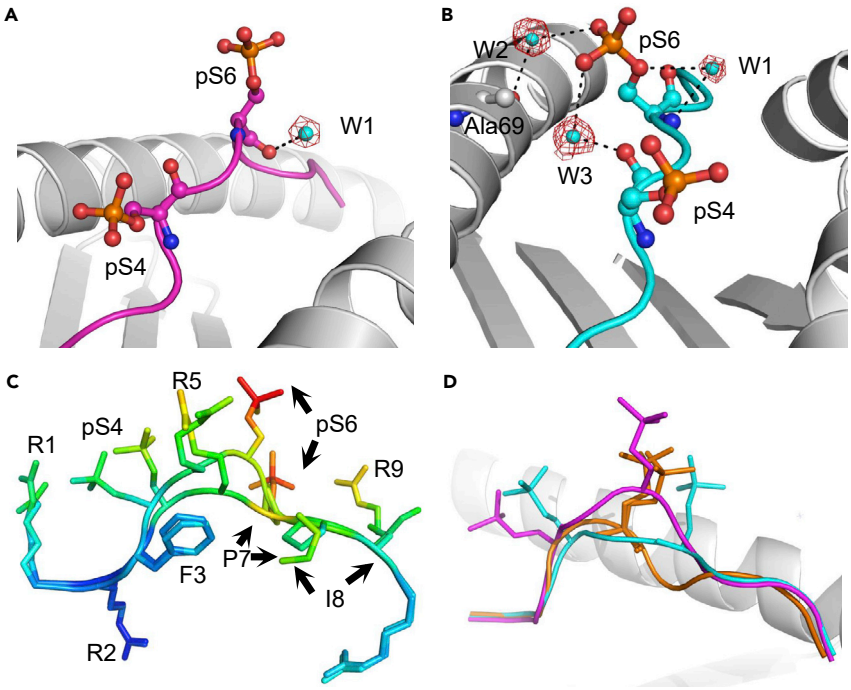


Figure 6. Unique structural characteristics of the P6 phosphosite and the solvent exposure

(A) In the structure of HLA-B*2705 complexed to the diphosphorylated 10-mer peptide, Dip-RR10, P6-Ser of the diphosphopeptide, Dip-RR10, displays complete solvent exposure and does not directly interact with the α1–α2 helices. Only a main chain contacting a water molecule (W1) was observed.

(B) In the HLA-B*2705 and diphosphorylated 9-mer Dip-RR9 complex, P6-pSer forms hydrogen bonds with water molecules to interact with the α1 helix (W2) and the peptide backbone (W1 and W3). Peptide side chains are shown as sphere-and-stick models, the red spheres represent oxygen atoms, the blue spheres represent nitrogen atoms, cyan spheres represent water molecules, orange spheres represent phosphorus atoms, and black-dotted lines represent hydrogen-bonding interactions.

(C) The peptides before the α1 helices in the HLA-B*2705 structure is superimposed for comparison. The backbone and side chains of the peptides are labeled with different colors based on different isotropic temperature factors (B factors).

(D) Structural overlap of Dip-RR10 (purple) and Dip-RR9 (cyan) in HLA-B*2705 with the P5-Ser peptide (PDB code: 3FQU, orange) presented by and HLA-A2. See also Figure S4.

residue or water molecule bound to the side chain of the P6-pSer residue (Figure 6A). However, the bulged conformation at the P6-pSer residue in Dip-RR9 is not as prominent as observed in Dip-RR10. Regardless, it still lacks direct interaction with surrounding intramolecular residues. Instead, there are three water-mediated interactions that bridge the O1P atom of P6-pSer to the O atom in the hydroxyl group of Ala69, the O1P atom to the O atom in the hydroxyl group of P4-pSer, and the O atom in the hydroxyl group to the N atom in P6-pSer (Figure 6B). These three "water bridges" stabilize the position and conformation of phosphate groups in antigen molecules by joining the phosphorylated side chain of the Ser residue to the α1-helix of the MHC molecules with peptide backbone. By comparing the P4 and P5 phosphorylations, the phosphate group added onto the P6-pSer residue was found to have the highest solvent-exposed position (Figures 6C and 6D). Moreover, the superposition of glycosylated MHC I antigens (H-2K^{b} and H-2D^{b}-restricted) revealed that the protrusion degree of the P6 phosphate group of Dip-RR10 is even higher than that of the polysaccharide group/glycosyl group in the glycosylated peptide (Glithero et al., 1999; Speir et al., 1999). This results in a highest protrusion degree among the modifying groups in currently resolved naturally modified epitopes (Figure S4). Thus, with high levels of group exposure, the structures of diphosphoantigens presented on MHC I molecules reflect their potential dominant immunogenicity during immune recognition by T cells.

Site-specific effects of phosphate groups on the binding and stability of HLA-B27

To further understand the effects of diphosphorylation of peptide antigens on the peptide-binding capacity of MHC I molecules and the effects of different phosphosites on the stability of pMHC, we determined

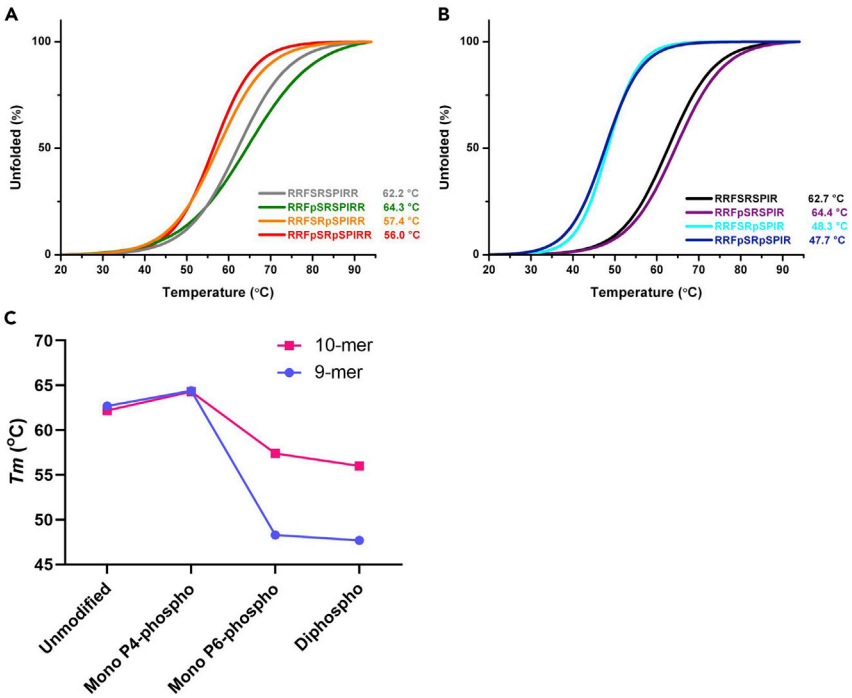


Figure 7. Impact of phosphorylated sites on the stability of HLA/peptide complexes

(A and B) Circular dichroism (CD) was used to measure the temperature changes of the HLA-B*2705 complex bound to unmodified, monophosphorylated P4-Ser or P6-Ser and diphosphorylated peptides on both P4-Ser and P6-Ser in the 10-mer RR10 (A) and its truncated 9-mer RR9 (B). The stability of the HLA-B*2705 complexes formed with peptides of different sequences are represented using different colored curves. The temperature (°C) values next to each sequence are the thermal denaturation values, T_ms.

(C) Polyline charts corresponding to the T_m values of different complexes in panels (A) and (B).

via CD spectroscopy the thermal stability of the HLA-B*2705 complexes with the peptides, including the above-mentioned four structurally resolved peptides and other monophosphorylated and nonmodified 10- or 9-residue peptides (Figures 7A–7C). Monophosphorylation of the P4-Ser residue enhances the binding affinity of both the 10- and 9-residue peptides with an increase of T_ms of 1.9 $\pm$ 0.2 °C. Unlike the P4 modification, there is a significant decrease in the binding affinity of both 10- or 9-residue peptides following P6 monophosphorylation. The significant decrease in binding affinity is likely attributable to loss of the role of the P6 residue as a secondary anchor with rotation of the peptide backbone following phosphorylation. The 9-residue peptide showed a more significant decrease in binding affinity than the 10-residue peptide. These data confirmed the important effects of the phosphosite-specific interactions observed in the crystal structure on the stabilization of the antigen-binding conformation.

DISCUSSION

In this study, we provided the first structural and thermodynamic insights into the presentation of multisite phosphopeptides by MHC I molecules. Previous structural and immunological studies on posttranslationally modified antigens (such as phosphorylated, acetylated, citrullinated, and glycosylated antigens) have been limited to mono-modifications (Beltrami et al., 2008; Petersen et al., 2009; Speir et al., 1999; Sun et al., 2014). Herein, we found that compared with the previously determined monophosphopeptides, the secondary phosphate group added to the peptide has a completely different conformation and intramolecular binding mode. This study has enhanced our understanding of the mechanism underlying the presentation of antigens that are modified at multiple sites.

At the molecular level, the majority of previously resolved structures of naturally modified MHC I antigens are characterized by the solvent exposure of the modification group, such as in glycosylation, N-terminal acetylation, and phosphorylation (Glithero et al., 1999; Mohammed et al., 2017; Sun et al., 2014). Such characteristics have also been observed in the diphosphopeptides resolved in this study. The phosphate

groups of both the P4 residue and P6 secondary anchor residue remain upward pointing outside the antigen-binding pockets. Perhaps, chemical groups added onto peptide antigens are more likely to become protruding sites on the surface of the antigens in immune responses, thus becoming recognition targets of the TCR. However, this mechanism likely relies on the large volume and chargeability of the modifying group itself. Our previous study indicated that the steric effects attributed to solvent exposure of N-terminal acetylation of peptide antigens by the flipping of carbon-carbon bonds (Sun et al., 2014). The secondary phosphate group can act as a direct potential target in TCR docking. Moreover, the addition of a secondary modifying group (such as P6 phosphorylation) triggers dramatic conformational changes at other sites in antigenic peptides, resulting in peptides with conformationally different antigenicity. Therefore, diphosphoantigens presented on MHC I molecules show conformationally variable antigenicity, and its high level of group exposure reflects the potential dominant immunogenicity of the phosphopeptide during immune surveillance by T cells.

In phosphorylation-dependent interactions, HLA-A2-restricted phosphopeptides consistently form a broad ionic bonding network with Arg65 and Lys66 residues on the α1-helix of the A2 molecule (Mohammed et al., 2008, 2017; Petersen et al., 2009). Although the Arg62 residue near the P4 phosphosite in the HLA-B*2705 molecule is also a positively charged amino acid residue, there is no direct interaction between the Arg62 residue and phosphate groups in the monophosphopeptide Monop-RR10 complex. However, such an interaction between Arg62 and the P4 phosphosite dramatically occurs in the Dip-RR10 complex owing to local conformational changes in the peptide triggered by the secondary phosphosite. Thus, the tendency of phosphate groups to form hydrogen bonds with the positively charged N-terminal and surrounding MHC residues relies on the sequence polymorphism of the MHC alleles.

Previous studies on the binding affinity of MHC I molecules to HLA-A2-restricted phosphoantigens using competition-based peptide binding and thermal stability assays confirmed that the phosphate group, which is a solvent-exposed group, enhances the peptide-binding affinity of MHC I molecules and the stability of the pMHC complex to a certain extent in a peptide-sequence-dependent manner (Mohammed et al., 2008; Petersen et al., 2009). This can be attributed to P4 phosphorylation and the broad ionic bonding network formed between the P1 residues and positively charged residues on the α1-helix of MHC molecules. However, we observed completely different thermochemical effects by the same type of modification at different sites on the same peptide antigen. The P4 phosphorylation in the HLA-B*2705-restricted peptide can have an HLA polymorphism-specific interaction with surrounding amino acid residues in HLA-B*2705, thereby enhancing the antigen-binding affinity and the stability of the pMHC as HLA-A2-restricted phosphopeptides (Petersen et al., 2009). Phosphorylation of the original anchor residue at position P6 results in loss of its role as an anchor residue owing to the steric hindrance of the phosphate group, leading to significant loss of its binding affinity to 10- or 9-residue antigens. Meanwhile, in the 10-mer diphosphorylated peptide, the phosphorylated P6 pushes the phosphorylated P4 to the P1-Arg anchor and makes the P4 phosphate group form the hydrogen bond network with the P1-Arg and also Arg62. In contrast, the 9-mer diphosphorylated peptide is stretched, and the phosphorylated P4 does not form the hydrogen bond network with the P1-Arg and also Arg62. So, the binding affinity of 9-mer diphosphorylated peptide is weaker than 10-mer diphosphorylated peptide.

Peptides diphosphorylated at an interval of two amino acid residues are the most common form of phosphopeptides (Schweiger and Linial, 2010), which was also substantiated in our study. Meanwhile, the phosphopeptides with a positively charged N-terminal residue and P4 phosphorylation (P5 in a small number of peptides) are common among the currently known phosphoantigens of MHC I molecules (Schweiger and Linial, 2010). In our study, the profile analysis of the protein phosphorylation site shows the high proportion of Arg as the third amino acid before the P1P phosphorylation form, indicating a potential RXXpSXpSXXX motif for diphosphopeptides. Thus, P4/P6 diphosphopeptides in this study are representative resolved structures of MHC I-restricted T cell peptides with this motif. The structures indicated that the impact of modification on the recognition of diphosphopeptides by T cells not only depends on peptide sequences and phosphate groups but also on the phosphosites. Nevertheless, the diphosphopeptides with no interval (PP), two intervals (P2P), or more interval residues between two phosphorylations may also act as the peptides presented on MHC I. Our structure investigations of the diphosphopeptides with the P1P mode (residue P4 and P6) showed a common feature of the diphosphopeptide presentation, when the biphosphorylation sites such as PP (residue P4 and P5) or P2P (residue P4 and P7) located in the middle region of the peptides. But the structural basis for the presentation of these diphosphopeptides still need further investigations.

In view of the fact that phosphorylation is an important marker for various physiological activities, such as carcinogenesis, malignant transformation, and apoptosis, there has been increasing attention on the application of phosphoantigens in immunotherapeutic strategies against tumors. The structural basis of the multisite phosphorylated epitope presentation and immune recognition may be important for the development of vaccines or pharmaceutical ingredients for the treatment of autoimmune diseases and cancer.

Limitations of the study

In our study, we mainly elucidate the motif of dual phosphorylated peptides and their presentation characteristics by MHC I through the structural analyses and human protein phosphorylation site database analyses. As for the functional T cell recognition of the peptides, because the phosphopeptides in this study are naturally processed phosphorylated peptides eluted from human cell expressed MHC I, they are hard to elicit a T cell response *in vivo*, and this result has been mentioned in the previous research (Zarling et al., 2000).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Microbe strains
- METHOD DETAILS
  - The phosphorylation sites
  - Peptide synthesis and preparation of expression constructs
  - Expression and purification of the HLA-B*2705 complexes
  - Crystallization and data collection
  - Structure determination and analysis
  - Thermostability measurements using circular dichroism (CD) spectroscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104013.

ACKNOWLEDGMENTS

This work was supported by grants from National Natural Science Foundation of China (81971501) and the National Key R&D Program of China (2021YFC0863400). W.J.L. is supported by the Excellent Young Scientist Program of the National Natural Science Foundation of China (81822040).

AUTHOR CONTRIBUTIONS

W.J.L. and G.F.G. conceived and designed the experiments. M.S. and L.F. performed the experiments. M.S., Y.Z., N.Z., X.L., C.Y., S.J., X.L., C.C.L.W., J.Q., and W.J.L. analyzed the data. Y.Z., M.S., C.C.L.W., W.J.L., and G.F.G. contributed to the writing of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 30, 2022
Revised: February 14, 2022
Accepted: February 25, 2022
Published: April 15, 2022

REFERENCES

Beltrami, A., Rossmann, M., Fiorillo, M.T., Paladini, F., Sorrentino, R., Saenger, W., Kumar, P., Ziegler, A., and Uchanska-Ziegler, B. (2008). Citrullination-dependent differential presentation of a self-peptide by HLA-B27 subtypes. J. Biol. Chem. *283*, 27189–27199.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Cohen, P. (2000). The regulation of protein function by multisite phosphorylation–a 25 year update. Trends Biochem. Sci. *25*, 596–601.

Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Doyle, H.A., and Mamula, M.J. (2001). Post-translational protein modifications in antigen recognition and autoimmunity. Trends Immunol. *22*, 443–449.

Eggleton, P., Haigh, R., and Winyard, P.G. (2008). Consequence of neo-antigenicity of the 'altered self'. Rheumatology (Oxford) *47*, 567–571.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Engelhard, V.H., Altrich-Vanlith, M., Ostankovitch, M., and Zarling, A.L. (2006). Post-translational modifications of naturally processed MHC-binding epitopes. Curr. Opin. Immunol. *18*, 92–97.

Glithero, A., Tormo, J., Haurum, J.S., Arsequell, G., Valencia, G., Edwards, J., Springer, S., Townsend, A., Pao, Y.L., Wormald, M., et al. (1999). Crystal structures of two H-2Db/glycopeptide complexes suggest a molecular basis for CTL cross-reactivity. Immunity *10*, 63–74.

Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. *31*, 3320–3323.

Huang, L.C., Ross, K.E., Baffi, T.R., Drabkin, H., Kochut, K.J., Ruan, Z., D'Eustachio, P., McSkimming, D., Arighi, C., Chen, C., et al. (2018). Integrative annotation and knowledge discovery of kinase post-translational modifications and cancer-associated mutations through federated protein ontologies and resources. Sci. Rep. *8*, 6518.

Laskowski, R.A., Macarthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. Acta Crystallogr. D Struct. Biol. *75*, 861–877. https://doi.org/10.1107/S2059798319011471.

Liu, J., Sun, Y., Qi, J., Chu, F., Wu, H., Gao, F., Li, T., Yan, J., and Gao, G.F. (2010a). The membrane protein of severe acute respiratory syndrome coronavirus acts as a dominant immunogen revealed by a clustering region of novel functionally and structurally defined cytotoxic T-lymphocyte epitopes. J. Infect. Dis. *202*, 1171–1180.

Liu, J., Wu, P., Gao, F., Qi, J., Kawana-Tachikawa, A., Xie, J., Vavricka, C.J., Iwamoto, A., Li, T., and Gao, G.F. (2010b). Novel immunodominant peptide presentation strategy: a featured HLA-A*2402-restricted cytotoxic T-lymphocyte epitope stabilized by intrachain hydrogen bonds from severe acute respiratory syndrome coronavirus nucleocapsid protein. J. Virol. *84*, 11849–11857.

Lu, D., Liu, K., Zhang, D., Yue, C., Lu, Q., Cheng, H., Wang, L., Chai, Y., Qi, J., Wang, L.F., et al. (2019). Peptide presentation by bat MHC class I provides new insight into the antiviral immunity of bats. PLoS Biol. *17*, e3000436.

Mohammed, F., Cobbold, M., Zarling, A.L., Salim, M., Barrett-Wilt, G.A., Shabanowitz, J., Hunt, D.F., Engelhard, V.H., and Willcox, B.E. (2008). Phosphorylation-dependent interaction between antigenic peptides and MHC class I: a molecular basis for the presentation of transformed self. Nat. Immunol. *9*, 1236–1243.

Mohammed, F., Stones, D.H., Zarling, A.L., Willcox, C.R., Shabanowitz, J., Cummings, K.L., Hunt, D.F., Cobbold, M., Engelhard, V.H., and Willcox, B.E. (2017). The antigenic identity of human class I MHC phosphopeptides is critically dependent upon phosphorylation status. Oncotarget *8*, 54160–54172.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Otwinoski, Z.M., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307.

Petersen, J., Wurzbacher, S.J., Williamson, N.A., Ramarathinam, S.H., Reid, H.H., Nair, A.K., Zhao, A.Y., Nastovska, R., Rudge, G., Rossjohn, J., et al. (2009). Phosphorylated self-peptides alter human leukocyte antigen class I-restricted antigen presentation and generate tumor-specific epitopes. Proc. Natl. Acad. Sci. U S A *106*, 2776–2781.

Schweiger, R., and Linial, M. (2010). Cooperativity within proximal phosphorylation sites is revealed from large-scale proteomics data. Biol. Direct *5*, 6.

Speir, J.A., Abdel-Motal, U.M., Jondal, M., and Wilson, I.A. (1999). Crystal structure of an MHC class I presented glycopeptide that generates carbohydrate-specific CTL. Immunity *10*, 51–61.

Sun, M., Liu, J., Qi, J., Tefsen, B., Shi, Y., Yan, J., and Gao, G.F. (2014). Nalpha-terminal acetylation for T cell recognition: molecular basis of MHC class I-restricted nalpha-acetylpeptide presentation. J. Immunol. *192*, 5509–5519.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *25*, 4876–4882.

Zarling, A.L., Ficarro, S.B., White, F.M., Shabanowitz, J., Hunt, D.F., and Engelhard, V.H. (2000). Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules *in vivo*. J. Exp. Med. *192*, 1755–1762.

Zarling, A.L., Polefrone, J.M., Evans, A.M., Mikesh, L.M., Shabanowitz, J., Lewis, S.T., Engelhard, V.H., and Hunt, D.F. (2006). Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. Proc. Natl. Acad. Sci. U S A *103*, 14889–14894.

Zhang, M., Li, H., He, Y., Sun, H., Xia, L., Wang, L., Sun, B., Ma, L., Zhang, G., Li, J., et al. (2015). Construction and deciphering of human phosphorylation-mediated signaling transduction networks. J. Proteome Res. *14*, 2745–2757.

Zhang, Q., Liu, K., Yue, C., Zhang, D., Lu, D., Xiao, W., Liu, P., Zhao, Y., Gao, G., Ding, C., et al. (2020). Strict assembly restriction of peptides from rabbit hemorrhagic disease virus presented by rabbit major histocompatibility complex class I molecule RLA-A1. J. Virol. *94*, e00396–20.

Zhang, S., Liu, J., Cheng, H., Tan, S., Qi, J., Yan, J., and Gao, G.F. (2011). Structural basis of cross-allele presentation by HLA-A*0301 and HLA-A*1101 revealed by two HIV-derived peptide complexes. Mol. Immunol. *49*, 395–401.

Zhu, S., Liu, K., Chai, Y., Wu, Y., Lu, D., Xiao, W., Cheng, H., Zhao, Y., Ding, C., Lyu, J., et al. (2019). Divergent peptide presentations of HLA-A*30 alleles revealed by structures with pathogen peptides. Front. Immunol. *10*, 1709.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli (*E. coli*) strain BL21 (DE3)	Novagen	Cat# 69450
Chemicals, peptides, and recombinant proteins		
Peptide RR10 (Son3, RRFSRSPIRR)	This paper	N/A
Peptide RR9 (RRFSRSPIR, the N-terminal truncated 9-mer of RR10)	This paper	N/A
Monophosphorylated peptides of RR10 (Monop-RR10, RRFpSRSPIRR)	Zarling et al., 2000	N/A
Monophosphorylated peptides of RR9 (Monop-RR9, RRFpSRSPIR)	This paper	N/A
Diphosphorylated peptides of RR10 (Dip-RR10, RRFpSRpSPIRR)	Zarling et al., 2000	N/A
Diphosphorylated peptides of RR9 (Dip-RR9, RRFpSRpSPIR)	Zarling et al., 2000	N/A
HLA-B*2705	This paper	N/A
human β_2-microglobulins	This paper	N/A
Critical commercial assays		
Crystal Screen Kit I/II and Index Kit	Hampton Research	http://www.hamptonresearch.com;
HiLoad 16/60 Superdex 200 pg column	GE Healthcare	Cat# 29194596
Deposited data		
HLA-B*2705/RR10 complex	This paper	PDB: 7CIQ
HLA-B*2705/Monop-RR10 complex	This paper	PDB: 7CIR
HLA-B*2705/Dip-RR10 complex	This paper	PDB: 7DYN
HLA-B*2705/Dip-RR9 complex	This paper	PDB: 7CIS
Recombinant DNA		
HLA-B*2705	This paper	N/A
human β_2-microglobulins	This paper	N/A
Software and algorithms		
Python version 2.7	Python Software Foundation	https://www.python.org
COOT	Emsley and Cowtan, 2004	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/; RRID: SCR_01422 2
Phenix	Liebschner et al., 2019	https://phenix-online.org/; RRID: SCR_01422 4
CCP4	Collaborative Computational Project	https://www.ccp4.ac.uk/ RRID: SCR_00725 5
PyMOL	Molecular Graphics System, Version 1.8 Schrödinger	https://pymol.org/2/ RRID: SCR_00030 5
Origin	OriginLab	https://www.originlab.com/ RRID: SCR_00281 5
Other		
PhoSigNet database	Zhang et all.,2015	http://119.3.70.71/PhoSigNet

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to lead contact, William J. Liu (liujun@ivdc.chinacdc.cn)

Materials availability

All unique reagents generated in this study are available from the lead contact without restriction.

Data and code availability

Atomic coordinates and structure factors have been deposited in the Protein Data Bank (http://www.rcsb.org) under accession codes 7CIQ, 7CIR, 7DYN and 7CIS for HLA-B*2705 in complex with RR10, Monop-RR10, Dip-RR10, and Dip-RR9, respectively. They are publicly available as of the date of publication. Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Microbe strains

BL21 (DE3) was cultured in LB with corresponding antibiotics after transfected with recombinant plasmid in 37°C incubator.

METHOD DETAILS

The phosphorylation sites

The phosphorylation sites information of 15705 human protein and related cancer information were acquired from PhoSigNet database (http://119.3.70.71/PhoSigNet/) (Zhang et al., 2015). According to the position information of the phosphorylation site of each protein in the database, the average number of different diphosphosite forms on tumor-association proteins and non-tumor-association proteins (average diphosphosite number/protein) are counted. The adjacent diphosphorylation were defined as PP, and the diphosphorylation forms with spaced amino acids were defined as PXP, etc (X = 1, 2, 3, ...). Then, the average number of diphosphorylation with one interval residues among tumor-association proteins (average P1P number/protein) in different diseases and the frequency of the third amino acid before the P1P phosphorylation form were analyzed.

Peptide synthesis and preparation of expression constructs

Peptides RR10 (Son3, RRFSRSPIRR) and RR9 (RRFSRSPIR, the N-terminal truncated 9-mer of RR10) were synthesized and purified to ~95% by reverse-phase high-performance liquid chromatography (HPLC) (Scilight-peptide, Inc. Beijing). The monophosphorylated peptides of RR10 (Monop-RR10, RRFpSRSPIRR) and RR9 (Monop-RR9, RRFpSRSPIR) indicated that have a phosphorylated Ser residue at the P4 positions, and diphosphorylated peptides of RR10 (Dip-RR10, RRFpSRpSPIRR) and RR9 (Dip-RR9, RRFpSRpSPIR) indiceted that have two phosphorylated Ser residues at the P4 and P6 positions. These phosphopeptides were also synthesized, and the phosphorylation modification was confirmed by mass spectrometry. The peptides were stored at −80 °C as a freeze-dried powder and dissolved in dimethyl sulfoxide (DMSO) before use. Furthermore, HLA-B*2705 cDNAs were synthesized (Genewiz Inc., Beijing). The expression plasmids of human β_2-microglobulins (β_2m, 1-99) were constructed in our laboratory (Zhang et al., 2011).

Expression and purification of the HLA-B*2705 complexes

Both HLA-B*2705 heavy chain and human β_2m proteins were expressed as inclusion bodies in *Escherichia coli* strain BL21 (DE3) and solubilized in 8 M urea buffer (Lu et al., 2019). The HLA-B*2705/peptide complexes were renatured from heavy-chain inclusion bodies and β_2m in the presence of peptides RR10, RR9, and their phosphorylated peptides by the dilution-refolding method (Zhang et al., 2020). Briefly, for dilution refolding, the HLA-B*2705 heavy chain, β_2m, and peptide in a 1:1:3 molar ratio was refolded by dilution in a 400-mM Arg·HCl solution. After 24 h of incubation at 277 K, the soluble proteins were concentrated, and the solution was changed to 20 mM Tris·HCl and 50 mM NaCl. The proteins were subsequently purified through a gel filtration HiLoad 16/60 Superdex 200 pg column (GE healthcare, Sweden) and centrifuged in a buffer of 20 mM Tris (pH 8.0) and 50 mM NaCl to a final concentration of 7–15 mg/mL for crystal screening.

Crystallization and data collection

Crystallization trials were performed via the hanging drop diffusion method using the Crystal Screen Kit I/II and Index Kit (Hampton Research). Based on experiences with the structural determination of HLA-B*2705 (Liu et al., 2010a, 2010b; Speir et al., 1999), we also performed direct screening using the gradient

conditions of 0.1 M Tris-HCl (pH 7.0–9.0) and 15–30% (w/v) PEG 8,000 buffer. The HLA-B*2705/Dip-RR10 crystals were obtained in a 0.2-M ammonium acetate, 0.1-M HEPES (pH 7.5), and 25% (w/v) PEG 3,350 solution at 291 K. Single crystals of HLA-B*2705/Dip-RR9 grew in 0.1 M Tris-HCl (pH 7.1) and 20.5% (w/v) PEG 8,000 at the same temperature. The crystals of HLA-B*2705/Monop-RR10 were generated under the conditions of 0.1 M Tris-HCl (pH 7.9) and 17.5% (w/v) PEG 8,000 at 291 K, and 0.2 M sodium chloride, whereas HLA-B*2705/RR10 was generated in 0.1 M Bis-Tris (pH 6.5), and 25% (w/v) PEG 3,350 at 277 K. Data from the HLA-B*2705/Dip-RR10 and HLA-B*2705/Dip-RR9 complexes were collected in-house on a Rigaku MicroMax007HF rotation anode X-ray generator at 40 kV and 20 mA (Cu Kα, λ = 1.5418 Å). Diffraction data for both HLA-B*2705/Monop-RR10 and HLA-B*2705/RR10 crystals were collected at 100 K by a KEK PF-AR NW12A (Japan) using a wavelength of 1.0 Å. In all cases, the crystals were first soaked in a reservoir solution containing 10–25% glycerol as a cryoprotectant for several seconds and the n flash-cooled in a stream of gaseous nitrogen at 100 K. The collected intensities were subsequently processed and scaled using the DENZO program and HKL2000 software package (HKL Research) (Otwinoski and Minor, 1997).

Structure determination and analysis

The structure of HLA-B*2705/RR10 was determined by molecular replacement using the determined HLA-B*2705 structure (Protein Data Bank [PDB] code 2BST) as a search model in the Crystallography & NMR System (CNS) program (Brunger et al., 1998). Clear solutions in both the rotation and translation functions corresponded to one molecule in one asymmetric unit. The peptide residues that differed in our structures compared to the search model were manually rebuilt in the COOT program under the guidance of F_o-F_c and $2F_o$-F_c electron density maps (Emsley and Cowtan, 2004). Subsequently, we refined the initial rigid body and performed a series of restrained translation, libration, and screw motion (TLS) refinements using the REFMAC5 program (Murshudov et al., 1997). Additional rounds of refinement were performed using the Phenix refine program implemented in the PHENIX package (Liebschner et al., 2019). We assessed the stereochemical quality of the final model using the PROCHECK program (Laskowski et al., 1993; Zhang et al., 2020). The structures of HLA-B*2705/Monop-RR10, HLA-B*2705/Dip-RR10, and HLA-B*2705/Dip-RR9 were determined using the final model of HLA-B*2705/RR10 through molecular replacement. The phosphate groups linked to serine in the phosphorylated peptides were also integrated to the peptides by insertion of the "SEP" group in COOT. The same refinement steps were performed to generate the final models. B factors were generated by utilizing the Baverage program in the CCP4i package (Collaborative Computational Project, Number 4, 1994). The molecular models and electron density-related figures were generated using PyMOL (http://www.pymol.org/). The sequence alignment was generated using Clustal X (Thompson et al., 1997) and ESPript (Gouet et al., 2003).

Thermostability measurements using circular dichroism (CD) spectroscopy

The thermostabilities of HLA-B*2705 complexed to different peptides were tested by using CD spectroscopy (Lu et al., 2019; Zhu et al., 2019). All complexes were refolded and purified beforehand, as described above, and measured at 0.2 mg/mL in a solution of 20 mM Tris (pH 8.0) and 50 mM NaCl. CD spectra at 218 nm were measured on a Chirascan spectrometer (Applied Photophysics) using a thermostatically controlled cuvette at temperature intervals of 0.2 °C and an ascending rate of 1 °C/min between 20 and 90 °C. The temperature of the sample solution was directly measured using a thermistor. The fraction of unfolded protein was calculated from the mean residue ellipticity. The unfolded fraction (%) is expressed as $(\theta-\theta_a) / (\theta_a-\theta_b)$, where θ_a and θ_b are the mean residue ellipticity values in the fully folded and fully unfolded states, respectively. The denaturation curves were generated by nonlinear fitting in OriginPro 8.0 (OriginLab). The midpoint transition temperature (T_m) was calculated by fitting data to the denaturation curves and using inflexion-determining derivatives.

QUANTIFICATION AND STATISTICAL ANALYSIS

This manuscript does not include quantification or statistical analysis.