Journal for ImmunoTherapy of Cancer

The landscape of MHC-presented phosphopeptides yields actionable shared tumor antigens for cancer immunotherapy across multiple HLA alleles

Zaki Molvi ,[1] Martin G Klatt ,[2,3,4] Tao Dao,[5] Jessica Urraca,[6] David A Scheinberg,[5,7] Richard J O'Reilly[1,6]

To cite: Molvi Z, Klatt MG, Dao T, *et al*. The landscape of MHC-presented phosphopeptides yields actionable shared tumor antigens for cancer immunotherapy across multiple HLA alleles. *Journal for ImmunoTherapy of Cancer* 2023;**11**:e006889. doi:10.1136/jitc-2023-006889

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/jitc-2023-006889).

Accepted 23 August 2023

For numbered affiliations see end of article.

Correspondence to
Dr Zaki Molvi;
zakimolvi@gmail.com

ABSTRACT

Background Certain phosphorylated peptides are differentially presented by major histocompatibility complex (MHC) molecules on cancer cells characterized by aberrant phosphorylation. Phosphopeptides presented in complex with the human leukocyte antigen HLA-A*02:01 provide a stability advantage over their non-phosphorylated counterparts. This stability is thought to contribute to enhanced immunogenicity. Whether tumor-associated phosphopeptides presented by other common alleles exhibit immunogenicity and structural characteristics similar to those presented by A*02:01 is unclear. Therefore, we determined the identity, structural features, and immunogenicity of phosphopeptides presented by the prevalent alleles HLA-A*03:01, HLA-A*11:01, HLA-C*07:01, and HLA-C*07:02.

Methods We isolated peptide-MHC complexes by immunoprecipitation from 11 healthy and neoplastic tissue samples using mass spectrometry, and then combined the resulting data with public immunopeptidomics data sets to assemble a curated set of phosphopeptides presented by 96 samples spanning 20 distinct healthy and neoplastic tissue types. We determined the biochemical features of selected phosphopeptides by in vitro binding assays and in silico docking, and their immunogenicity by analyzing healthy donor T cells for phosphopeptide-specific multimer binding and cytokine production.

Results We identified a subset of phosphopeptides presented by HLA-A*03:01, A*11:01, C*07:01 and C*07:02 on multiple tumor types, particularly lymphomas and leukemias, but not healthy tissues. These phosphopeptides are products of genes essential to lymphoma and leukemia survival. The presented phosphopeptides generally exhibited similar or worse binding to A*03:01 than their non-phosphorylated counterparts. HLA-C*07:01 generally presented phosphopeptides but not their unmodified counterparts. Phosphopeptide binding to HLA-C*07:01 was dependent on B-pocket interactions that were absent in HLA-C*07:02. While HLA-A*02:01 and HLA-A*11:01 phosphopeptide-specific T cells could be readily detected in an autologous setting even when the non-phosphorylated peptide was co-presented, HLA-A*03:01 or HLA-C*07:01 phosphopeptides were repeatedly non-immunogenic, requiring use of allogeneic T cells to induce phosphopeptide-specific T cells.

Conclusions Phosphopeptides presented by multiple alleles that are differentially expressed on tumors

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Phosphorylated peptides presented by the common human leukocyte antigen (HLA) alleles A*02:01 and B*07:02 are differentially expressed by multiple tumor types, exhibit structural fitness due to phosphorylation, and are targets of healthy donor T-cell surveillance, but it is not clear, however, whether such features apply to phosphopeptides presented by other common HLA alleles.

WHAT THIS STUDY ADDS

⇒ We investigated the tumor presentation, binding, structural features, and immunogenicity of phosphopeptides to the prevalent alleles A*03:01, A*11:01, C*07:01, and C*07:02, selected on the basis of their presentation by malignant cells but not normal cells. We found tumor antigens derived from genetic dependencies, that is, genes essential to the survival and proliferation of lymphomas and leukemias that bind HLA-A3, HLA-A11, HLA-C7 molecules. While we could detect circulating T-cell responses in healthy individuals to A*02:01 and A*11:01 phosphopeptides, we did not find such responses to A*03:01 or C*07:01 phosphopeptides, except when using allogeneic donor T cells, indicating that these phosphopeptides may not be immunogenic in an autologous setting but can still be targeted by other means.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ An expanded patient population expressing alleles other than A*02:01 can be addressed through the development of immunotherapies specific for phosphopeptides profiled in the present work, provided the nuances we describe between alleles are taken into consideration.

constitute tumor-specific antigens that could be targeted for cancer immunotherapy, but the immunogenicity of such phosphopeptides is not a general feature. In particular, phosphopeptides presented by HLA-A*02:01 and A*11:01 exhibit consistent immunogenicity, while phosphopeptides presented by HLA-A*03:01 and C*07:01, although appropriately presented, are not immunogenic. Thus, to address an expanded patient population, phosphopeptide-targeted immunotherapies should be wary of allele-specific differences.

BACKGROUND

T cells recognizing tumor-selective antigens presented in the context of major histocompatibility complex (MHC) are capable of inducing durable regressions of cancers refractory to standard treatment when adoptively transferred into patients. Adequate selection of antigen targets is critical to induce remission and prevent relapse. Neoantigens produced by non-synonymous somatic mutations represent a tumor-exclusive class of antigens but are typically private to each patient and more likely to be present when tumor mutational burden is sufficiently high. Differentially expressed tumor-associated antigens, such as WT1, Survivin, PRAME, and NYESO1 have been safely targeted to treat a variety of pediatric and hematologic tumors,[1–3] which generally harbor too few mutations to produce adequate neoantigens. A significant pitfall in antigen selection is insufficient peptide presentation in that tumor antigen epitopes that are found to be immunogenic are not necessarily endogenously presented by tumors. Advances in mass spectrometry (MS) solve this problem by enabling direct sequencing of the immunopeptidome by identifying peptides eluted from human leukocyte antigen (HLA) immunoprecipitates (HLA-IP).[4] HLA-IP followed by MS of cell lines and primary samples has enabled identification of post-translationally modified peptides, such as phosphopeptides[5] and glycopeptides[6] as an emerging class of tumor antigens.

Several phosphopeptides have been shown to be presented by certain HLA class I and II alleles that exhibit enhanced immunogenicity, a feature hypothesized to be due to unique structural features.[7–9] Interestingly, Cobbold *et al* have shown that while healthy donors harbor T-cell responses specific to these phosphopeptides, patients with leukemia lack such responses.[10] Moreover, these responses are restored post-allogeneic stem cell transplant. Patients with colorectal cancer have also been found to harbor tumor infiltrating lymphocytes that recognize phosphopeptides, as well as peripheral T cells that recognize phosphopeptides at higher frequencies than healthy donors,[11] closely mimicking studies of neoantigens produced by somatic mutations.[12 13]

We sought to expand the known landscape of phosphopeptides presented by prevalent HLA molecules that could be useful as immunotherapy targets. We applied HLA-IP to 10 hematologic cell line samples. We then combined our data with public immunopeptidomics data sets to assemble an expanded data set of phosphopeptides not presented by normal tissues. From this data set, we selected phosphopeptides of interest presented by HLA-A*03:01, HLA-A*11:01, and HLA-C*07:01 and studied their binding stability in vitro, in silico structural properties, and ability to stimulate T cells, to identify candidates for future immunotherapeutic treatment strategies.

METHODS

Cells

T2 cells, Epstein-Barr virus (EBV)-transformed lymphoblastoid B cells (BLCL), and monoclonal EBV-associated lymphoproliferative disease cells (EBV-LPD) emerging post-marrow allograft were maintained in RPMI 1640 Medium +10% fetal bovine serum, supplemented with 2 mM L-glutamine, and penicillin-streptomycin. T2 cells expressing HLA alleles were generated as follows: HLA-A0301, HLA-C0701, and HLA-C0702 encoding complementary DNA (cDNA) (Integrated DNA Technologies) were inserted into pSBbi-GP[14] (Addgene plasmid # 60511) using the NEBuilder HiFi Assembly Master Mix. Successful ligation was confirmed by Sanger sequencing of DH5a (NEB) colonies transformed with the ligation reaction. Stable transfection was performed by transfecting T2 cells with 4.5 µg of HLA cDNA in the pSBbi-GP backbone and 0.5 µg of pCMV(CAT)T7-SB100[15] (Addgene plasmid # 34879), followed by puromycin selection.

HLA ligand identification

HLA ligands were isolated and identified using immunoprecipitation and liquid chromatography-tandem mass spectrometry (LC-MS/MS) as described previously.[16] Approximately 1–2×10^8 cells were washed in phosphate buffered saline (PBS), snap-frozen, and stored at −80°C. Pellets were thawed on ice and lysed in 1% 3-([3-cholamidopropyl] dimethylammonio)−1-propanesulfonate in PBS supplemented with Roche cOmplete protease inhibitor cocktail and Roche PhosSTOP for 1 hour at 4°C. Lysates were cleared by centrifugation. Supernatants were circulated over W6/32-conjugated sepharose columns for MHC class I isolation and either L243-conjugated or IVA12-conjugated sepharose columns for MHC class II isolation using a peristaltic pump overnight at 4°C. Peptide-HLA complexes were eluted from dried columns in 1% trifluoroacetic acid (TFA). Peptide-HLA complexes were adsorbed onto Sep-Pak tC18 columns (Waters) pre-equilibrated with 80% acetonitrile (ACN). Peptides were eluted in 40% ACN 0.1% TFA or 30% ACN 0.1% TFA. Solid-phase extraction of peptide eluates was performed using in-house C18 minicolumns (Empore) washed with 80% ACN/0.1% TFA and pre-equilibrated with 1% TFA. Peptide eluates were run through the C18 minicolumn, which was subsequently washed twice with 1% TFA, and desalted peptides were eluted with 80% ACN/0.1% TFA. Samples were analyzed using a Lumos Fusion operated in data-dependent acquisition (DDA) mode. Peptides were separated using a 12 cm built-in-emitter column using a 70 min gradient (2–30%B, B: 80% ACN/0.1%

formic acid). 3 µL of 8 µL were injected. 3+ and 4+ (and undetermined charge states) peptides were allowed in the mass range: m/z 250–700. 2+ peptides were selected in the mass range m/z 350–1000 while 1+ peptides were selected in the range: m/z 750–1800. We have made our MS search results available in online supplemental file 1. We have included NetMHCpan4.0 predictions for all peptides on which we performed binding predictions in online supplemental file 2.

T2 stabilization

T2 cells expressing the indicated HLA molecule were harvested from culture and incubated for 18 hours at room temperature. Cells were subsequently washed with PBS and resuspended in serum-free RPMI with 3 µg/mL beta-2-microglobulin (MP Biomedicals) and the indicated concentration of peptide for 3 hours at room temperature followed by 3 hours at 37°C in 5% CO_2. Cells were washed, stained for 30 min at 4°C with Fixable Viability Dye Violet (1:1,000), HLA-A2 PE (BD, 1:100), and HLA-A3 PE-Vio770 (Miltenyi, 1:100) or HLA-C AF647 (BioLegend, 1:100). Stained cells were washed twice and analyzed on a BD LSR II.

Molecular docking

Peptide docking to HLA molecules was performed as described previously.[17] Solved crystal structures of peptide-HLA complexes were retrieved from the Protein Data Bank (PDB) and used as template structures. PDB entry 5VGE was used to dock 9-mers to HLA-C0702, and 3RL1 and 3RL2 were used for docking 9-mers and 10-mers, respectively, to HLA-A0301. To generate a template for HLA-C0701, UCSF Chimera[18] was used to incorporate the K66N and S99Y mutations that distinguish C0701 from C0702. Peptides of interest were threaded onto the template by mutating the peptide using the Dunbrack and SwissSidechain rotamer libraries[19 20] implemented in UCSF Chimera. Structures were prepacked and docked using the FlexPepDock protocol in refinement mode in Rosetta3.[21 22] For each distinct peptide-HLA complex, each complex was scored in Rosetta energy units (REU) using the Rosetta3 full-atom score function REF2015, and the top 10 lowest REU models were selected among 200 high-resolution models. UCSF Chimera was used to visualize models and analyze hydrogen bonds.

Immunogenicity assessment

Phosphopeptides were assessed for their immunogenicity by ELISpot analysis or multimer staining of sensitized HLA-typed donor peripheral blood mononuclear cells (PBMC). PBMC were sensitized to phosphopeptides selected for each donor's HLA typing using a method described previously.[4] For dendritic cell (DC) priming, the method of Wölfl and Greenberg[23] was used. In some cases, autologous peptide-pulsed CD14+ cells or T2 cells expressing the relevant HLA allele, were used instead of DC. On days 10–13 after initial sensitization or priming, cells were restimulated with peptide-pulsed, lethally irradiated autologous PBMC, autologous DC, or T2 cells and thereafter maintained in media containing interleukin (IL)-7/15 at 5 ng/mL and IL-2 (Miltenyi) at 50 IU/mL. All cultures were maintained in X-VIVO 15 media (Lonza)+5% human AB serum (Gemini). Cells were typically analyzed on days 10–13 of each stimulation cycle via multimer staining or ELISpot against autologous peptide-pulsed targets as described previously.[24–26]

Dextramers were assembled according to our protocol,[27] based on a previous method,[28] for HLA-A*11:01, A*02:01, A*03:01, and C*07:01 because for these HLA alleles we could obtain commercially available UV-exchange MHC monomers (BioLegend). For dextramer enrichment from PBMC, 1–3×10^8 PBMC were incubated in 50 nM dasatinib and FcR block for 30 min at 37°C, then incubated with 10 µg/mL of the indicated dextramer pool for 1 hour at 4°C, and enriched with anti-Cy5, anti-PE, or anti-Cy7 beads (Miltenyi) sequentially over two LS Columns (Miltenyi). Enriched cells were either stained for surface markers and analyzed by flow cytometry or expanded in 96-well plates with irradiated allogeneic feeders and anti-CD3/28 reagent (STEMCELL) in the presence of IL-2 300 IU/mL, IL-7 5 ng/mL, and IL-15 5 ng/mL for 10–14 days until subsequent analysis or further enrichment and re-expansion. When dextramer-enriched samples were acquired by flow cytometry, the entire sample volume was acquired.

Data analysis and statistics

Flow cytometry data was analyzed in FCS Express (De Novo Software). All other data were analyzed using custom Python and R scripts. All statistical analyses presented used a paired t-test. For genetic dependency analysis, data were downloaded from DepMap portal[29] (www.depmap.org/). For each group of cell lines, DEMETER2 scores were averaged for each gene using a custom Python script. Gene Ontology (GO) analysis was performed using STRING (www.string-db.org). Peptide predictions were performed using NetMHCpan4.0 or NetMHCphosPan1.0. Peptide motif analysis was performed using GibbsCluster-2.0. MS data from previous studies[4 30–32] were obtained on ProteomeXchange (accession nos: PXD004746, PXD005704, PXD012083, PXD013831). Kinase enrichment analysis was performed using KEA3.[33]

RESULTS

Identification of immunogenic phosphopeptides presented in the hematologic malignant state

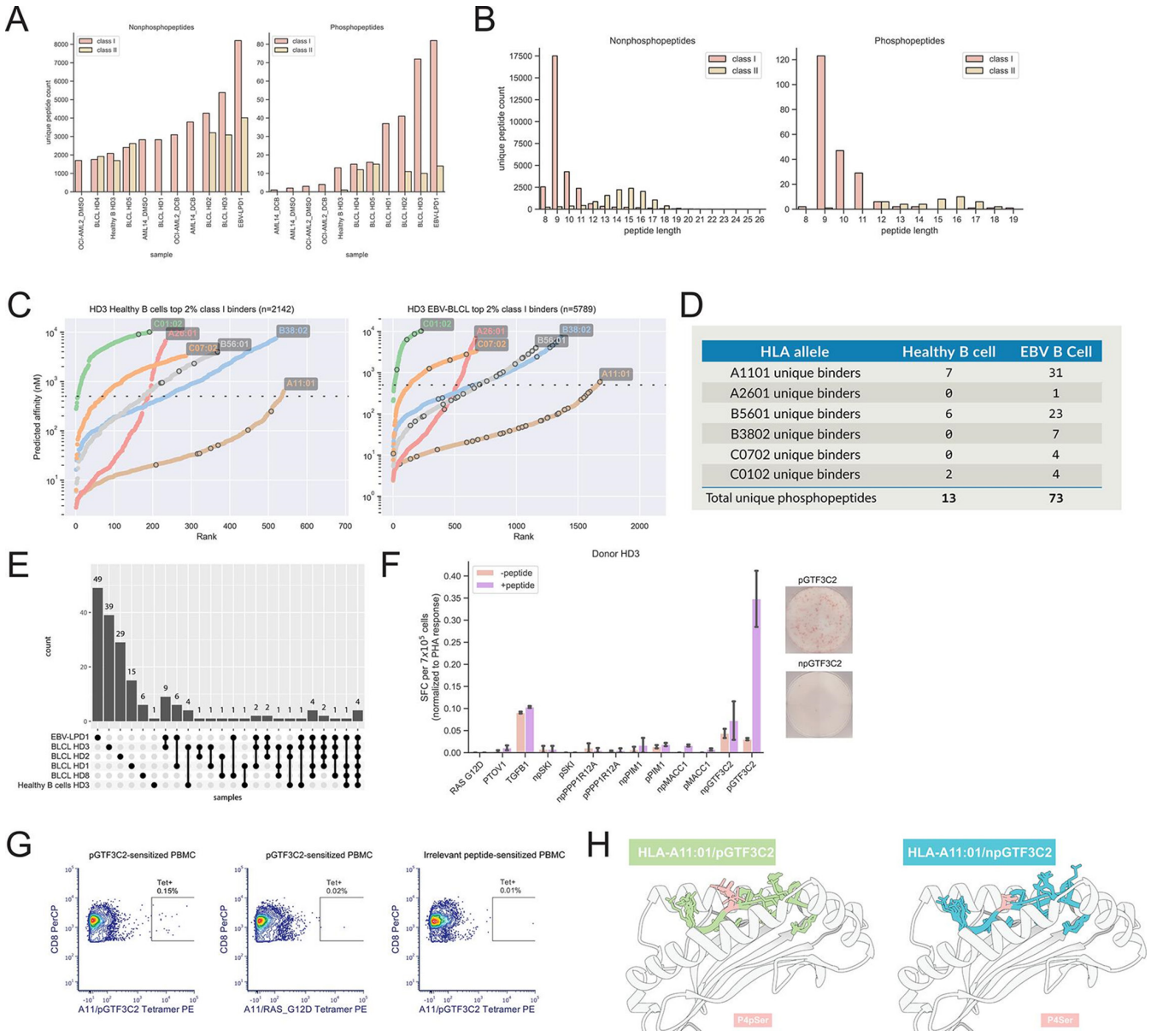
We attempted to elucidate the phosphopeptidome for common HLA alleles other than A*02:01 to define tumor antigens presented by malignant cells, but not by normal tissue. Because of the documented presentation of phosphopeptides on EBV-BLCL, we used HLA class I and II-based immunoprecipitation and LC-MS/MS to isolate HLA ligands from six EBV-BLCL lines as well as two acute myeloid leukemia (AML) lines treated with either decitabine to enhance antigen presentation

or dimethylsulfoxide (DMSO), one EBV-LPD line, and one healthy B-cell sample. We focused our sample selection preferentially on HLA-A*11:01+ samples because our initial studies identified HLA-A*11:01 as a prevalent allele from which we could elute and confidently identify a substantial set of phosphopeptides. By restricting the resultant peptide identifications to >8-mer peptides filtered by a stringent false discovery rate (FDR) of 1% and DeltaMod >20 in the Byonic environment,[34] we recovered 40,557 unique non-phosphopeptides and 255 unique phosphopeptides across all samples. HLA class I peptidomes contained 214 unique phosphopeptides derived from 194 human proteins, whereas class II peptidomes contained 53 unique phosphopeptides derived from 37 proteins. EBV-transformed samples, such as EBV-BLCL and EBV-LPD, were the top-ranking samples when comparing samples by counts of unique unmodified peptides and phosphopeptides in our data set (figure 1A). HLA class I and II phosphopeptides were most frequently 9-mers and 16-mers (figure 1B), respectively, in agreement with previous studies.[8 32] To elucidate the differences in phosphopeptidome of B cells in the malignant versus healthy state in an autologous setting, we examined class I peptides and phosphopeptides eluted from equal numbers of EBV-BLCL and healthy B cells from the same donor (donor HD3). When constrained to peptides with a NetMHCpan4-predicted percentile rank, which only considers each peptide's primary sequence and is agnostic to phosphorylation, of <2% for each donor-expressed allele, high affinity peptides were more numerous for the HLA-A*11:01 allele than other allotypes present in Donor HD3 (figure 1C). We found more than twice the number of predicted high affinity peptides presented by donor HD3's EBV-BLCL compared with their autologous healthy B cells (5,789 vs 2,142 peptides; figure 1C). The observed difference in peptidomes between these two types of B cells could be due to differences in MHC expression between cultured EBV-BLCL and ex vivo-enriched B cells, as well as due to loss from the stringent 1% FDR used to analyze samples. Nonetheless, phosphopeptides, demarcated by gray circles in figure 1C, were presented in higher numbers by each allele in EBV-BLCL compared with autologous healthy B cells (enumerated in figure 1D; data in online supplemental file 2). On the basis of their amino acid sequences excluding phosphorylation, most of these phosphopeptides were assigned to HLA-A*11:01 by NetMHCpan4.0. Ten phosphopeptides shared between both EBV-BLCL and autologous healthy B cells are listed in online supplemental table 2. By analyzing the overlaps in phosphopeptidomes of all six HLA-A*11:01+ samples in our data set (figure 1E), we found previously undescribed phosphopeptides pGTF3C2, pPPP1R12A, pPIM1, pMYBBP1A, and pSRRM1 were presented by multiple EBV-BLCL and EBV-LPD, but not healthy B cells.

Since some phosphopeptides have been shown to be immunogenic in A2+ donors,[10 17 35] we sought to determine if A11+ donors also harbor phosphopeptide-specific T cells (PP-CTL). We selected phosphopeptides and normal peptides presented by donor HD3's EBV-BLCL on the basis of their high predicted binding affinity and limited presentation on normal cells, and found that only pGTF3C2 elicited peptide-specific interferon (IFN)-g production by sensitized T cells from this donor (figure 1F). Similar to previous reports, the unmodified peptide, npGTF3C2, did not induce specific T cells. These results were corroborated by tetramer staining of donor HD3's PBMC after a single stimulation with pGTF3C2 (figure 1G). To explain the difference in immunogenicity, we used a molecular docking approach that we previously applied to explain the specificity of a T cell receptor (TCR)-mimicking antibody recognizing the A2/pIRS2 phosphopeptide.[17] We docked pGTF3C2 and npGTF3C2 to HLA-A*11:01 using FlexPepDock,[21] and examined the top 10 lowest energy models for each complex (figure 1H). Both models exhibited similar mean energy scores (−746 REU for pGTF3C2 vs −739 REU for npGTF3C2) and peptide backbone conformations (figure 1H), consistent with the observation that npGTF3C2 is predicted to be a strong binder, exhibiting a 0.19% percentile rank for binding to HLA-A*11:01, and that npGTF3C2 was co-presented with pGTF3C2 in our MS data (online supplemental file 1). These results demonstrated that certain phosphopeptides, such as pGTF3C2, presented by HLA-A*11:01 on EBV-BLCL but not healthy autologous B cells, can mobilize PP-CTL responses in a normal donor even when the unmodified peptide is co-presented by HLA-A*11:01.

An expanded phosphopeptide data set yields shared HLA-A3/A11 phosphopeptide tumor antigens

As shown in figure 2A, we further expanded our data set to include phosphopeptides curated from previous studies,[4 30–32] yielding a data set containing 2,466 distinct phosphopeptides from 96 sample data sets spanning 20 tissue types, predominantly AML (n=19), mantle cell lymphoma (MCL) (n=19), meningioma (n=18), and EBV-BLCL (n=13) (samples listed in online supplemental file 3). To extend our studies of the A*11:01 allele, we focused on phosphopeptides presented by HLA-A*03:01, which shares a common binding motif with A*11:01. We found 772 phosphopeptides presented by A3+ or A11+ samples but not by samples expressing other alleles and compared their presentation across all samples to find recurringly presented phosphopeptides (figure 2B; online supplemental file 4). Hierarchical clustering of the samples presenting these phosphopeptides did not necessarily correlate with their tissue of origin or cancer type; however, such an analysis is limited by the uneven distribution of tissues and HLA alleles in the data set. Unsupervised clustering of the 719≥9-mer A3/A11 phosphopeptide sequences revealed two motifs: 68% belong to a motif dominated by P4 Ser and P5 Pro, a preference consistently reported for phosphopeptides[4 10 32]; 32% belong to a motif characterized by repeated leucines, reflecting the representation of A*03:01 samples in the



HLA allele	Healthy B cell	EBV B Cell
A1101 unique binders	7	31
A2601 unique binders	0	1
B5601 unique binders	6	23
B3802 unique binders	0	7
C0702 unique binders	0	4
C0102 unique binders	2	4
Total unique phosphopeptides	**13**	**73**

Figure 1 The A*11:01 immunopeptidome contains recurringly presented, immunogenic phosphopeptides. (A) Unique peptide counts for class I and II non-phosphopeptides (left) and phosphopeptides (right) across all samples on which we performed HLA-IP. (B) Length distribution of class I and II non-phosphopeptides (left) and phosphopeptides (right). (C) NetMHCpan4.0 predicted affinity versus peptide affinity rank for each HLA allotype present in healthy B cells (left) and EBV-BLCL (right) from the same donor (Donor HD3). Gray circles indicate the position of phosphopeptides. Dashed line indicates 500 nM affinity. (D) Total counts of unique phosphopeptides between Donor HD3's autologous healthy B cells and EBV-BLCL visualized in C. (E) UpSet plot illustrating the distribution and intersection of unique phosphopeptides among HLA-A*11:01+ samples on which we performed HLA-IP. The plot comprises a bar chart, with each bar representing a count of unique phosphopeptides that are mutually found in the samples indicated with black circles below, and its height denoting the count of distinct phosphopeptides shared among samples in the intersection. (F) ELISpot of Donor HD3 PBMC sensitized to the indicated peptides, expressed as fraction of PHA-stimulated control for each peptide-sensitized culture. Each peptide-sensitized co-culture was split into three groups of two replicates on the ELISpot plate: an unpulsed peptide autologous APC (–peptide), an autologous APC pulsed with the relevant peptide (+peptide), and PHA. The spot-forming cells (SFC) for –peptide and +peptide wells were each divided by the mean PHA responses to provide an appropriate comparison of specific interferon-g response across different peptide-sensitized groups. Representative images for pGTF3C2-sensitized and npGTF3C2-sensitized PBMC are shown to the right. (G) Tetramer staining of Donor HD3 PBMC showing increased frequency of pGTF3C2 tetramer-specific CD8 T cells after pGTF3C2 sensitization relative to both irrelevant tetramer (A11/RAS_G12D) stained cells and irrelevant peptide-sensitized autologous PBMC. (H) Docking results of the top 10 lowest energy models of pGTF3C2 (left) or npGTF3C2 (right) in complex with HLA-A*11:01, with p4Ser highlighted. AML, acute myeloid leukemia; DCB, decitabine; EBV-BLCL, Epstein-Barr virus transformed B lymphoblastoid cell line; EBV-LPD, Epstein-Barr virus-associated B lymphoproliferative disease; HD, healthy donor; HLA, human leukocyte antigen; HLA-IP, HLA immunoprecipitation; PBMC, peripheral blood mononuclear cell; APC, allophycocyanin; PHA, Phytohaemagglutinin; DMSO, dimethylsulfoxide.

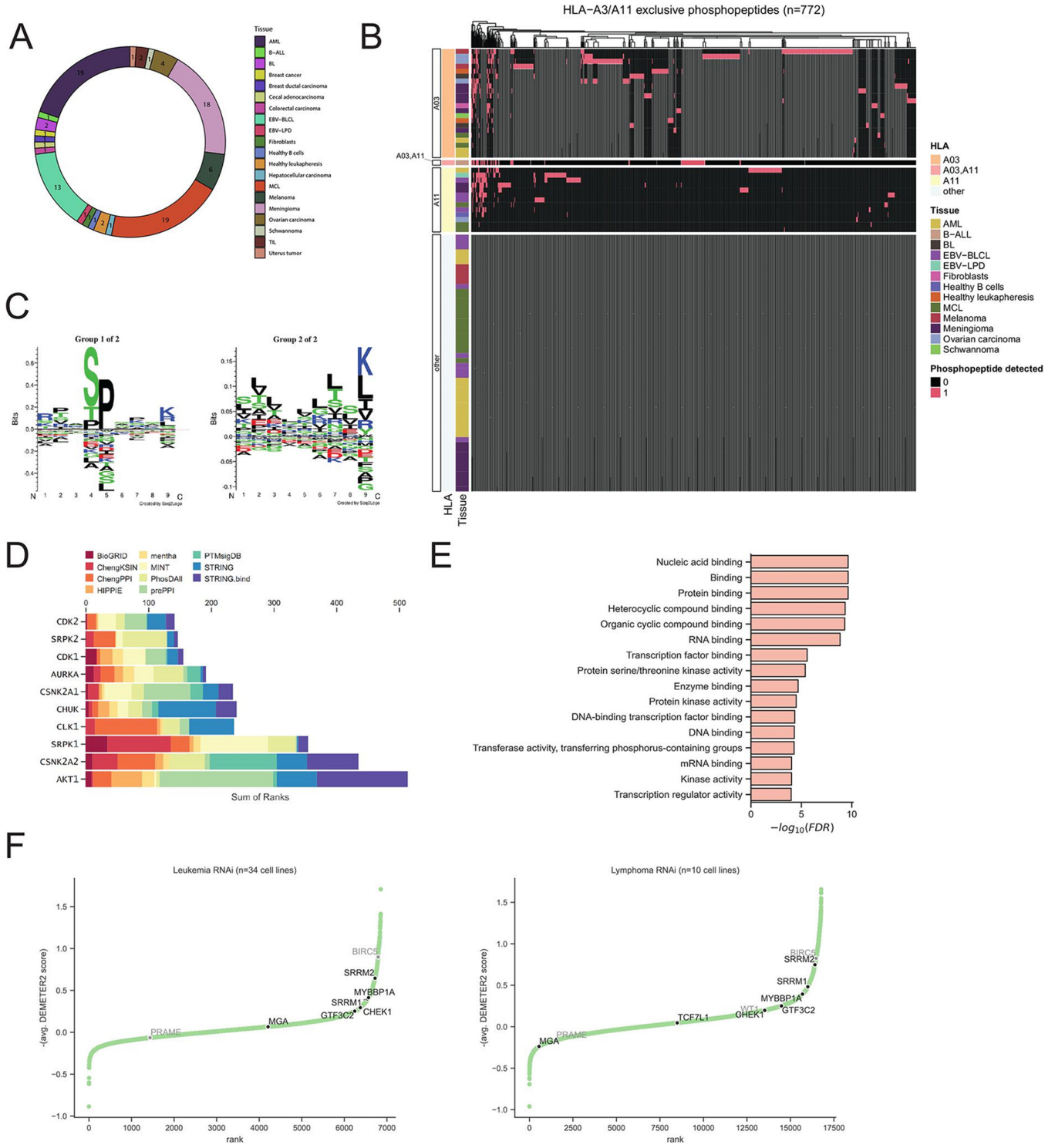


Figure 2 Expanded data set of phosphopeptides. (A) Circle plot of tissue types represented in an expanded data set showing counts of unique samples of each tissue type. (B) Heatmap visualization of HLA-A3/A11 phosphopeptide presentation. Vertical axis represents distinct samples. Horizontal axis represents a distinct phosphopeptide. Samples are annotated on the left by HLA allotype and tissue type. (C) Motif analysis of phosphopeptides in B. (D) Kinases inferred to be upstream of parental genes of the phosphopeptides visualized in B. Kinases are ranked by their MeanRank, with MeanRank decreasing from bottom to top, and plotted against the sum of their ranks in each of the kinase libraries, which are indicated by color. (E) Gene Ontology term analysis of parental genes of phosphopeptides in (B). (F) Rank plots of genetic dependency score (−1*average (DEMETER2 score)) for lymphoma and leukemia cell lines from pooled RNAi screens[29] with parental genes of phosphopeptides denoted, summarized in table 1. Genetic dependency increases from bottom to top. Known tumor-associated antigens Survivin (BIRC5) and WT1 are denoted in gray for reference. AML, acute myeloid leukemia; B-ALL, B-cell acute lymphoid leukemia; EBV-BLCL, Epstein-Barr virus transformed B lymphoblastoid cell line; EBV-LPD, EBV-associated B lymphoproliferative disease; HLA, human leukocyte antigen; MCL, mantle cell lymphoma; mRNA, messenger RNA; FDR, False Discovery Rate.

Table 1 Summary of selected phosphopeptides from the expanded data set presented by HLA-A3 and/or HLA-A11 most frequently presented in samples visualized figure 2B and further analyzed in figure 2F

Sequence	No. malignant samples presenting phosphopeptide (no. healthy samples presenting phosphopeptide)	Gene	HLA	Malignant tissue presentation
RVAsPTSGVK	15 (3)	IRS2	A03, A11	B-ALL, melanoma, ovarian carcinoma, meningioma, schwannoma, MCL
SVSsPVKSK*	15 (1)	MGA	A03, A11	EBV-LPD, EBV-BLCL, B-ALL, melanoma
RTNsPGFQK	12 (0)	RBM26	A03, A11	EBV-LPD, EBV-BLCL, B-ALL, ovarian carcinoma, meningioma, BL
HVYtPSTTK	11 (3)	ANKRA2	A03, A11	B-ALL, melanoma, meningioma, schwannoma, AML
RTAsPPPPPK*	11 (0)	SRRM1	A03, A11	EBV-LPD, EBV-BLCL, melanoma, AML, MCL, BL
KLRsPFLQK	10 (2)	DBNL	A03, A11	B-ALL, meningioma, AML, BL
KVQGsPLKK	10 (1)	AKAP12	A03, A11	B-ALL, ovarian carcinoma, meningioma, schwannoma
KVSsPTKPK*	10 (1)	GTF3C2	A03, A11	EBV-LPD, EBV-BLCL, B-ALL, melanoma, ovarian carcinoma
RAKsPISLK	10 (1)	CARD11	A03, A11	EBV-BLCL, meningioma, schwannoma, MCL
RLSsPISKR	10 (1)	BARD1	A03, A11	Melanoma, ovarian carcinoma, meningioma, AML, BL
ATAsPPRQK	9 (1)	SRRM2	A11	EBV-LPD, EBV-BLCL, B-ALL, meningioma
ATQsPISKK*	9 (0)	MYBBP1A	A03, A11	EBV-BLCL, EBV-LPD, B-ALL, ovarian carcinoma, meningioma
GSGsPAPPR	9 (1)	GPATCH8	A11	EBV-LPD, EBV-BLCL, B-ALL, meningioma
SVKsPVTVK	9 (1)	TCF7L1	A03, A11	Meningioma, melanoma, schwannoma
RTAsPNRAGK*	3 (0)	HIF1A	A03, A11	EBV-LPD, B-ALL, melanoma
RTAsPPALPK*	4 (0)	PRDM2	A03, A11	EBV-BLCL, EBV-LPD, melanoma, BL
HSLsPGPSK*	4 (0)	PIM1	A11	EBV-BLCL, meningioma
ATPTSPIKK*	8 (0)	PPP1R12A	A11	EBV-BLCL, EBV-LPD, B-ALL, meningioma, MCL

*Denotes phosphopeptides selected for further study.
AML, acute myeloid leukemia; B-ALL, B-cell acute lymphoid leukemia; BL, Burkitt's lymphoma; EBV-BLCL, Epstein-Barr virus transformed B lymphoblastoid cell line; EBV-LPD, EBV-associated B lymphoproliferative disease; HLA, human leukocyte antigen; MCL, mantle cell lymphoma.

data set since submotifs of A*03:01 contain an over-representation of leucine[36] (figure 2C). The most recurringly presented A3/A11 phosphopeptides in our data set are summarized in table 1. The majority of these phosphopeptides were detected on both A3+ and A11+ samples. Several phosphopeptides in table 1, such as those derived from MGA, RBM26, and SRRM1 are presented on EBV-BLCL in addition to other lymphoid malignancies, such as B-cell acute lymphoid leukemia, MCL, and AML. In particular, the phosphopeptide pSRRM1 (RTAsPPPPPK) was presented on 11 transformed cells, hematologic or solid tumors and absent in the healthy tissues in our data. Moreover, its primary sequence is not detected in any cadaver tissues profiled in the HLA ligand atlas[37] which only contains unmodified peptides. The absence of the phosphopeptide in our healthy tissue data sets and its corresponding unmodified peptide in non-malignant cadaver tissues suggests limited expression of this epitope in healthy tissues, therefore representing a potentially broad immunotherapeutic target.

To infer the upstream kinase pathways involved in phosphorylation, we performed kinase enrichment analysis[33] of the 254 parental genes for all phosphopeptides presented by at least two malignant A3/A11 samples, but not by healthy samples. This analysis ranked the essential kinase CDK2 as the most probable upstream kinase,

which is involved in proliferation and DNA repair. Kinases are plotted by decreasing MeanRank versus sum of ranks in each kinase library in figure 2D. However, concluding that this limited set of kinases are responsible for phosphorylating these phosphopeptides requires further experimentation. We also examined Clinical Proteomic Tumor Analysis Consortium (CPTAC) data to determine if phosphatase copy numbers influenced phosphorylation levels of IRS2 Ser1100, the phosphosite of the well-studied pIRS2 phosphopeptide (table 1), but found the samples size too limiting to make adequate conclusions (online supplemental figure 1).[38] GO analysis revealed that the phosphopeptides visualized in figure 2B preferentially derive from nucleic-acid binding proteins (figure 2E). Moreover, an exploratory analysis revealed that certain HLA-A3/A11 phosphopeptides shown in table 1 that were recurringly presented were derived from genes that have been shown to be essential to proliferation of leukemia and lymphoma cells in genome-wide RNAi screens,[39] such as SRRM1, SRRM2, GTF3C2, and MYBBP1A (figure 2F). That the HLA-A3/A11 phosphopeptidome samples peptides from genes essential to lymphoma and leukemia cell proliferation supports the categorization of selected phosphopeptides as shared tumor antigens.

Structural features of phosphopeptide complexes

Phopshopeptide-MHC complexes have been postulated to be immunogenic by virtue of their phosphate moiety conferring an increase in both MHC stabilization and solvent-facing character.[9 40] However, this characteristic is not a universal feature as structural studies of phosphopeptide-HLA-A2 complexes have shown this feature varies on a peptide-by-peptide basis.[9 41] To determine if phosphorylation conferred augmented peptide binding to HLA-A3 molecules, we determined the binding of selected phosphopeptides and their unmodified counterparts to MHC via in vitro and in silico assays. Using transporter associated with antigen processing (TAP)-deficient T2 cells, modified to express HLA-A*03:01, we found that in five of five phosphopeptide-unmodified pairs, both peptides could stabilize HLA-A*03:01 with no promiscuous binding to HLA-A*02:01, but the phosphorylated peptide did not confer increased stabilization of HLA-A*03:01 compared with its unmodified counterpart (figure 3A–C). In the case of one HLA-A3 phosphopeptide, pMGAP, phosphorylation appears to dampen the capacity of the peptide to stabilize HLA-A3, requiring a higher concentration of peptide to achieve the same degree of stabilization as unmodified MGAP, similar to what has been reported for phosphopeptides binding B*07:02 and B*40:01.[32 42] To illustrate more precisely how phosphorylation could negatively impact the stability of a phosphopeptide-MHC complex, we performed an example analysis by docking the MGAP phosphopeptide pMGAP and its unmodified counterpart npMGAP to HLA-A3 to determine the differences conferred by phosphorylation. By examining the top 10 lowest energy models for pMGAP and npMGAP, we found they exhibit nearly congruent backbone conformations (figure 3D). The mean number of hydrogen bonds formed at the binding interface was not significantly different (11.7 for npMGAP vs 12.1 for pMGAP; two-tailed p value=0.39), leading us to further investigate the interfacial interactions. Because the Rosetta score is a weighted linear combination of individual energy terms, we decomposed the Rosetta energy function into its individual terms and calculated the mean $\Delta\Delta G$ of each term between the npMGAP and pMGAP models to clarify the energetic favorability underlying npMGAP. The strongest energetically favorable change was observed in electrostatic potential (Δfa_elec, −54.5 kcal/mol), which was offset by an unfavorable but smaller magnitude change in solvation energy (Δfa_sol) of 29.5 kcal/mol (figure 3E). Since the only difference between the pMGAP and npMGAP structures is the presence or absence, respectively, of a phosphate moiety on P4Ser, we examined which residues on HLA interact with P4Ser to stabilize npMGAP, finding that in npMGAP, P4Ser interacts with Asn66 with significantly stronger potentials (range: −0.869 to −0.096 kcal/mol, mean: −0.574 kcal/mol), than those of P4pSer with Asn66 in pMGAP (range: −0.091 to −0.004 kcal/mol, mean: −0.0346 kcal/mol) (figure 3F). Visualizing P4Ser in relation to Asn66 reveals sufficient van der Waals radii[43] overlap to constitute seven contacts between these two residues in the npMGAP/HLA-A3 structure (figure 3G, left); however, in the pMGAP/HLA-A3 structure, P4Ser does not achieve sufficient proximity to Asn66 to make such contacts (figure 3G, right), presumably to avoid steric clashes. Due to the disparate behavior of individual phosphopeptide-MHC complexes, it cannot be concluded that this interaction explains the relatively weaker affinity of the A3 phosphopeptides we tested. Since docking energies for other phosphopeptide/HLA-A3 complexes were not concordant with observed in vitro binding, we limit our conclusions to the case of pMGAP/HLA-A3. Nevertheless, since 3/5 phosphopeptide-unmodified pairs that we examined were co-presented in our MS data, our results demonstrate that HLA-A3 can capably present both phosphopeptides and their unmodified counterparts.

Given the limited polymorphism of the HLA-C locus relative to the HLA-A and HLA-B loci, we sought to characterize phosphopeptides presented by prevalent HLA-C alleles represented in our data set for which we could assemble phosphopeptide dextramers as potential tumor antigens. To this end, we selected phosphopeptides found on multiple HLA-C*0701 and HLA-C*0702 expressing EBV-BLCL, tumors, and monoallelic cell lines,[44] excluding phosphopeptides whose sequences were detected in healthy cadaver tissue.[37] We selected seven phosphopeptides, summarized in online supplemental table 5, from pRBM14, pRAF1, pMYO9B, pZNF518A, pWNK, pSTMN, and pNCOR, and pulsed them onto T2 cells, modified to express HLA-C*0701 or HLA-C*0702 to measure their capacity to stabilize each HLA. Excepting pWNK, all of the selected phosphopeptides were found on both C0701+ and C0702+ cell line immunopeptidomes. In

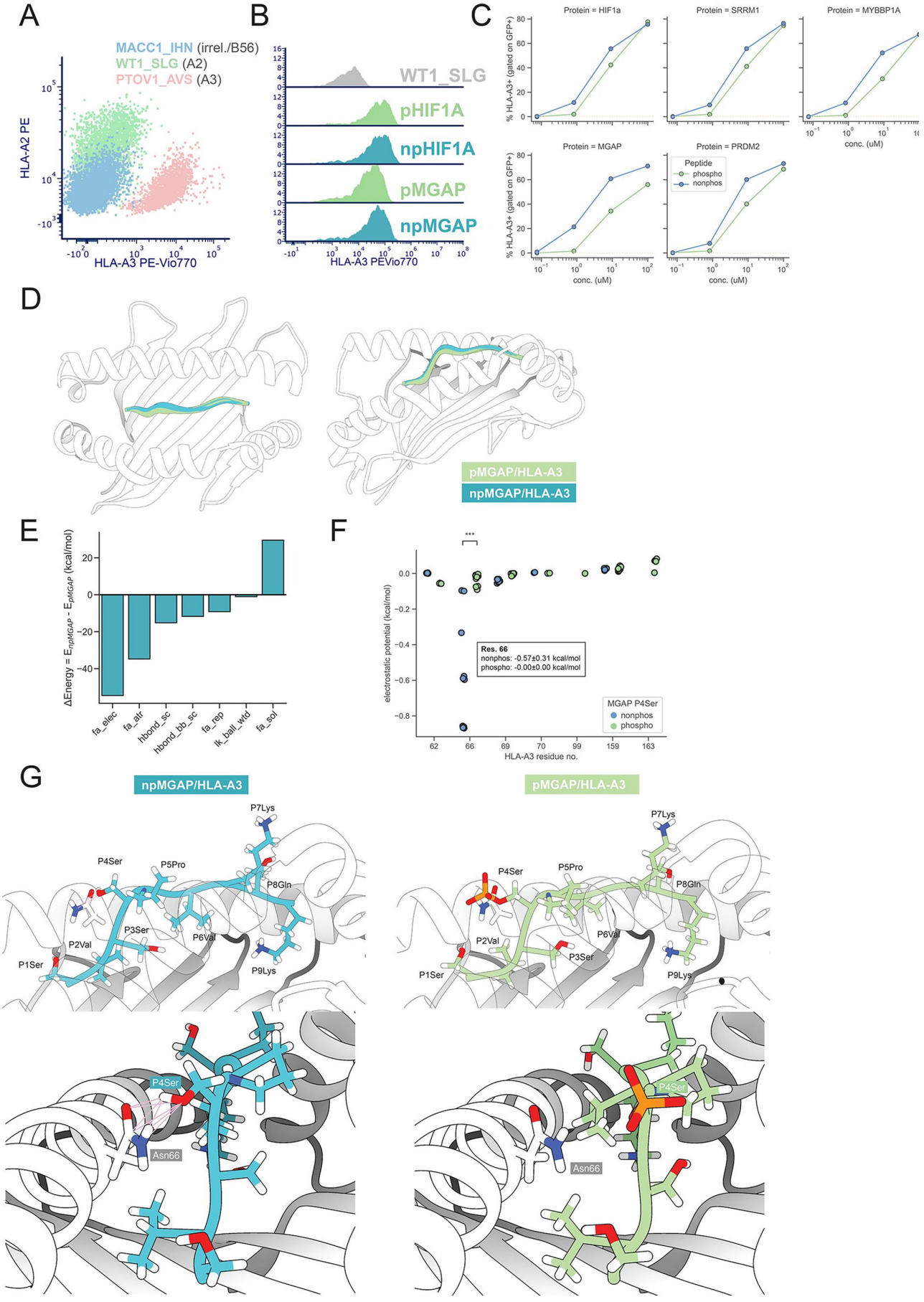


Figure 3 (Continued)

Figure 3 HLA-A0301 phosphopeptide binding properties. (A) Validation of HLA-A0301 stabilization using reference HLA-A3-binding peptide (PTOV1), A2-binding peptide (WT1_SLG), and irrelevant peptide (MACC1). (B) Histograms of HLA-A3 staining on T2-A0301 pulsed with 100 µg/mL of the indication peptides. (C) HLA-A3 stabilization in response to dose titration of phosphopeptide-non-phosphopeptide pairs. (D) Visualization of peptide backbone conformations of top 10 models for pMGAP and npMGAP in complex with HLA-A3. (E) Decomposition of mean $\Delta\Delta G$ (ΔEnergy) by Rosetta energy score terms between the top 10 npMGAP and pMGAP models visualized in (D). (F) Electrostatic potential of all P4Ser interactions with HLA-A3 for pMGAP and npMGAP. The mean±SD values for residue 66 are shown in the box. (G) Visualization of P4Ser interacting with Asn66 in MGAPwt/HLA-A3 (left) and pMGAP/HLA-A3 (right) complexes. Pink lines indicate van der Waals overlaps sufficient to produce contacts between P4Ser and Asn66. All statistics produced by or paired (F) t-test. P value annotation legend: ns, 5.00e−02< p ≤1.00e+00; *: 1.00e−02< p ≤5.00e−02; **: 1.00e−03< p ≤1.00e−02; ***: 1.00e−04< p ≤1.00e−03; ****: p ≤1.00e−04. HLA, human leukocyte antigen.

spite of this, none of them stabilized C*0702, but all of them stabilized C*0701 on T2 cells (figure 4A), despite comparable baseline expression levels for both alleles on T2 cells (online supplemental figure 3). To investigate the molecular determinants of binding to HLA-C7, we docked each 9-mer phosphopeptide onto either C*0701 or C*0702. For each phosphopeptide complex, the top 10 models were consistently more stable for C*0701 than C*0702 phosphopeptide complexes, as measured by a lower Rosetta energy score (figure 4B). Despite the significant energetic difference between the complexes, the peptide backbone conformations were similar between C*0701 and C*0702 complexes, with C*0701-bound phosphopeptides being slightly more shifted out of the HLA groove (figure 4C); however, the buried interfacial solvent accessible surface area was only significantly different between pNCOR-bound C*0701 and C*0702 complexes (figure 4D). Examining representative structures for pNCOR shows that the discrepancies of conformations between C7 molecules is mediated by different sidechain orientations (figure 4E). Since C*0701 and C*0702 differ by only two residues at positions 66 and 99, we hypothesized that direct contacts with these residues may explain the difference in conformations and interfacial surface area. pNCOR makes 12 and 13 hydrogen bonds to C*0702 and C*0701, 10 of these hydrogen bonds are mutual to both complexes (online supplemental table 4). When pNCOR is bound to C*0701, the sidechain of P2Arg makes two hydrogen bonds with Tyr99 and one with Asp9, both belonging to the β sheet of the HLA B-pocket. When bound to C*0702, which bears a serine at position 99 instead of tyrosine, P2Arg did not form a hydrogen bond with the distant Ser99 molecule (figure 4F), thereby losing two hydrogen bonds in the B-pocket β sheet. This change is associated with C*0702-bound pNCOR forming additional hydrogen bonds with the HLA α1 helix between the sidechain of P1Arg and Glu63, and between the phosphate moiety of P4pSer and Arg69. The loss of these hydrogen bonds in the C*0702 phosphopeptide complex is associated with a peptide conformation that is more buried in the HLA groove and less energetically favorable. These results suggest altered B-pocket interactions underlie the stability of phosphopeptides when bound to C*0701. The fact that 5/7 selected phosphopeptides differentially expressed by tumors bound C0701 provides further rationale for targeting these phosphopeptides in an expanded patient population.

Immunogenicity assessment

The immunogenicity of HLA-A*02:01-presented phosphopeptides has previously been shown to be mediated by pre-existing memory T cells in healthy donors.[10 35] This distinguishes phosphopeptides from other self-antigens which are typically recognized by lower frequency naïve T cells. Given the prevalence of PP-CTL, we sought to detect PP-CTL via phosphopeptide dextramers assembled using an improved method.[28] We assembled HLA-A*0201 dextramers in complex with six selected phosphopeptides as well as an irrelevant A2-binding WT1 peptide, WT1_SLG, and used fluorophore barcoding[45] to discriminate phosphopeptide specificity from WT1 specificity in figure 5A, with predicted binding in online supplemental table 3. We stimulated A*02:01$^+$ donor PBMC with A*02:01$^+$ T2 cells pulsed with the selected six phosphopeptides and found that after only 10 days, 4% phosphopeptide dextramer$^+$ CD8$^+$ T cells could be detected that did not cross-react with WT1_SLG dextramer. However, PBMC from the same donor contemporaneously stimulated with T2 cells pulsed with the A2-binding WT1_SLG peptide did not yield T cells specific for either WT1_SLG or phosphopeptides (figure 5B), providing evidence that the immunogenicity of these phosphopeptides may be more pronounced than that of WT1_SLG despite our observation that these phosphopeptides and WT1_SLG comparably stabilize HLA-A2. In two additional A*02:01+ healthy donor buffy coats, we were able to use sequential magnetic enrichment to achieve a highly pure (>95% dextramer$^+$) PP-CTL population (figure 5C). To determine if PP-CTL could exert effector function, we measured the specific IFN-g response of PP-CTL by ELISpot. Across four A*02:01+ donors, responses could be detected to pIRS2 and pCDC25B (figure 5D and E), the latter of which were consistently phosphorylation-specific.

Since we observed responses to certain A*11:01 phosphopeptides, we assembled an A*11:01 dextramer panel encoding either phosphopeptide or irrelevant RAS G12V specificity (figure 5F). When these dextramers were used to enrich and expand PBMC from an A*11:01+ healthy donor for PP-CTL, a small population of PP-CTL could be observed that did not cross-react to irrelevant RAS G12V peptide dextramers (figure 5G). Unlike our

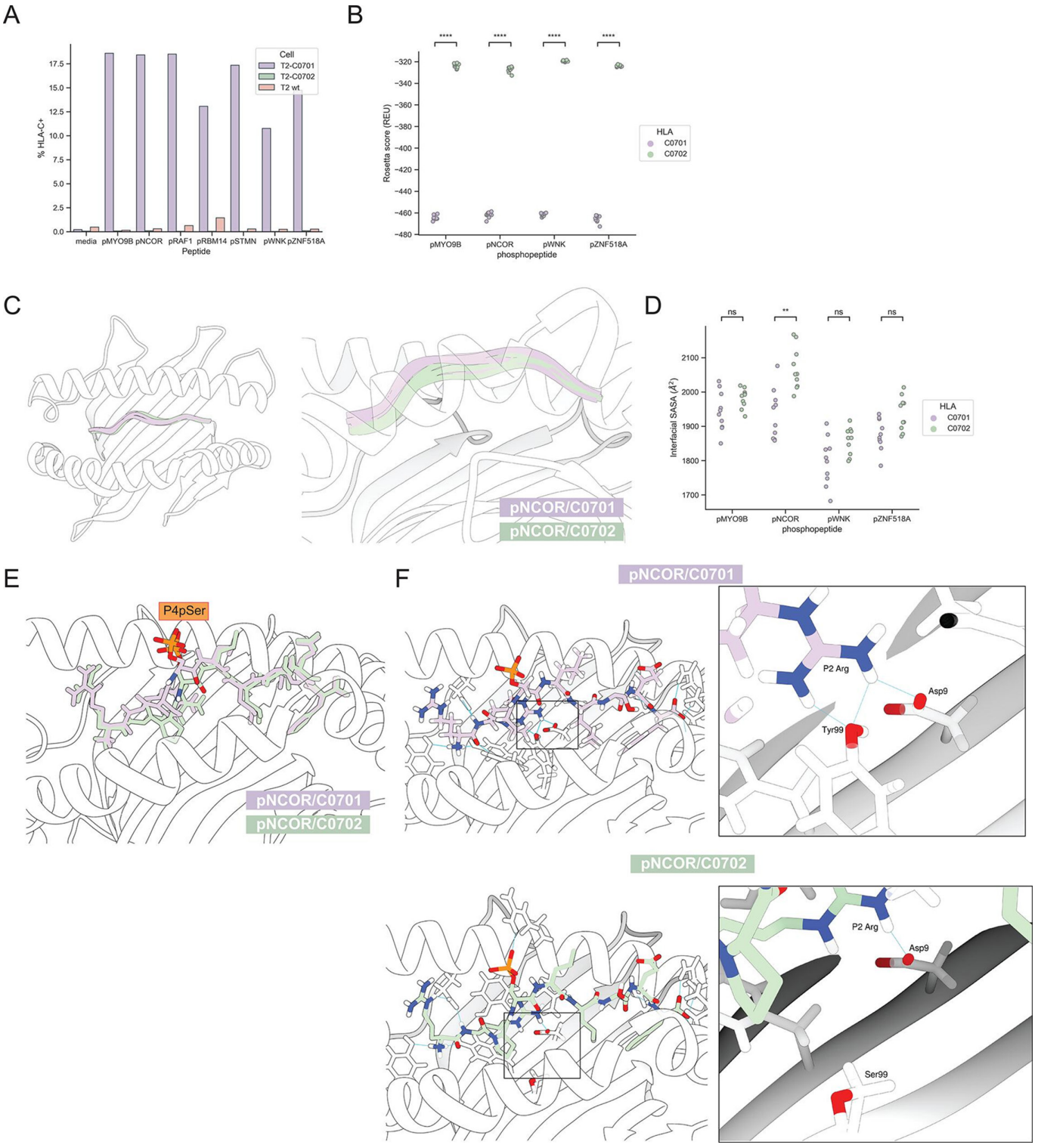


Figure 4 HLA-C7 phosphopeptide binding properties. (A) HLA-C stabilization of indicated phosphopeptides at 100 µg/mL for T2-C0701, T2-C0702 or wt. (B) Rosetta score of top 10 models produced by FlexPepDock for each indicated phosphopeptide in complex with HLA-C0701 and HLA-C0702. (C) Visualization of peptide backbone conformation for the top 10 models for each of pNCOR/C0701 and pNCOR/C0702. (D) Buried interfacial solvent accessible surface area (SASA) at the interface between each indicated phosphopeptide and C0701 or C0702. (E) Representative visualization of sidechains for pNCOR/C0701 and pNCOR/C0702. (F) Visualization of hydrogen bonds formed between P2Arg and Tyr99 in the pNCOR/C0701 complex (top) and absence of hydrogen bonds between P2Arg and Ser99 in pNCOR/C0702 (bottom). All statistics produced by paired t-test. P value annotation legend: ns, 5.00e−02< p ≤1.00e+00; *: 1.00e−02< p ≤5.00e−02; **: 1.00e−03< p ≤1.00e−02; ***: 1.00e−04< p ≤1.00e−03; ****: p ≤1.00e−04. HLA, human leukocyte antigen.

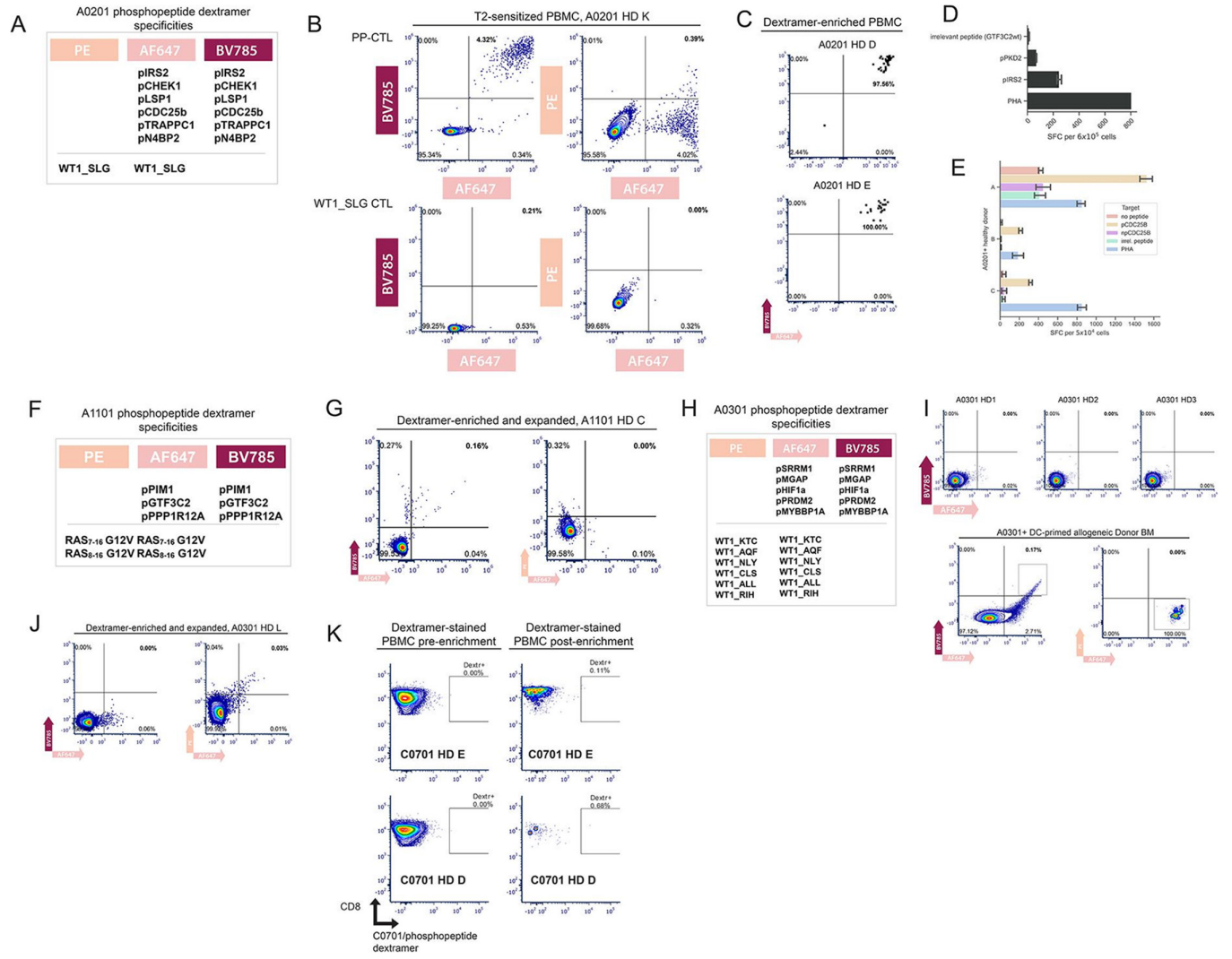


Figure 5 Combinatorial fluorophore barcoded dextramer analysis of PP-CTL. (A) HLA-A0201 dextramer panel. HLA-A2/phosphopeptide dextramer is encoded by AF647/BV785 and A2/WT1_SLG is encoded by PE/AF647 combination. (B) A0201+ donor T cells sensitized to either phosphopeptides (top) or WT1_SLG (bottom). PP-CTL (top) can be detected by A2/phosphopeptide dextramers in the AF647/BV785 channel and do not cross-react to A2/WT1_SLG dextramer in the PE/AF647 channel. (C) Analysis of A0201+ donor buffy coats enriched for PP-CTL by sequential dextramer enrichment over two magnetic columns and then analyzed by flow cytometry. (D) ELISpot of A0201 healthy donor T cells sensitized to A2 phosphopeptides pPKD2 and pIRS2. PHA response is shown as positive control and irrelevant npGTF3C2 peptide response serves as negative control. (E) ELISpot of three A0201 donors (denoted A, B, C on y-axis) whose T cells were repeatedly sensitized to pCDC25B and then rechallenged with autologous CD14+ cells pulsed with either of the indicated peptides listed under "Target". PHA serves as a positive control. (F) A1101 dextramer panel encoding phosphopeptides on BV785/AF647 and RAS G12V peptides on PE/AF647. (G) A1101 Donor PBMC enriched with dextramers and expanded show binding to A1101/phosphopeptide dextramers (AF647/BV785) but not irrelevant RAS G12V dextramers (PE/AF647). (H) HLA-A0301 dextramer panel encoding phosphopeptides on AF647/BV785 multiple A3-binding WT1 epitopes on PE/AF647, used to assess responses in panels I, J. (I) Three A0301+ donors primed to the phosphopeptides in H. Did not result in expansion of A3/phosphopeptide dextramer-binding cells (top), but A0301– negative donor BMMC stimulated with A0301+ DC produce A3/phosphopeptide dextramer-binding T cells that do not cross-react with A3/WT1 dextramer. (J) Dextramer-enrichment and expansion increases frequency of A3/WT1-specific T cells, but not PP-CTL. (K) C0701 Healthy donors do not have directly observable PP-CTL after dextramer enrichment with C0701/phosphopeptide dextramers on a single fluorochrome. DC, dendritic cell; HD, healthy donor; HLA, human leukocyte antigen; PBMC, peripheral blood mononuclear cell; PP-CTL, phosphopeptide-specific T cells; SFC, spot-forming cells; PHA, phytohaemagglutinin; BMMC, bone marrow mononuclear cells.

observations of A*02:01 PP-CTL, we did not observe A*11:01 PP-CTL directly after sequential magnetic enrichment of PBMC from two additional A*11:01+ healthy donor buffy coats (data not shown). Since our enrichment scheme co-enriches for RAS G12V specificity, our data suggest that A*11:01 phosphopeptide responses are at least more prevalent than those of RAS G12V, but still not as prominent as responses to A*02:01 phosphopeptides.

To detect HLA-A*03:01 PP-CTL, we assembled A*03:01 dextramers using shared phosphopeptides identified from

our analysis denoted in table 1 and used A3-binding epitopes of WT1 as irrelevant controls (figure 5H). We could not detect PP-CTL directly ex vivo after dextramer enrichment of A*03:01 PBMC (data not shown), nor could we observe PP-CTL after priming with phosphopeptide-loaded autologous DCs in three A*03:01+ healthy donors (figure 5I, top) or repeated sensitization with phosphopeptide-pulsed T2-A0301 (data not shown). However, after priming A*03:01-negative donor BMMC-derived T cells with allogeneic phosphopeptide-loaded A0301+ DC, we could observe PP-CTL that bound A3 phosphopeptide dextramer (figure 5I, bottom left), but not irrelevant A3 WT1 dextramer (figure 5I, bottom right). While the recognition of HLA-A3 by these allogeneic T cells appears to be dependent on the presence of phosphopeptides, other HLA alleles may still be promiscuously recognized by such an allogeneic population. Interestingly, when we used a co-enrichment scheme specific for the shared WT1/phosphopeptide fluorochrome AF647 and expanded the resulting T cells in an A*03:01+ donor (figure 5J), a minor population of WT1-specific (figure 5J, right panel), rather than phosphopeptide-specific (figure 5, left panel), T cells was observed. Overall, the A*03:01 phosphopeptides we selected appear to be less immunogenic than A*03:01 WT1 peptides, and A*11:01 and A*02:01 phosphopeptides, requiring, in our hands, stimulation of A*03:01-negative allogeneic T cells with phosphopeptide-loaded A*03:01 targets to elicit a phosphopeptide-specific response.

To determine if there were any immediate responses to phosphopeptides binding C*07:01 that we previously described, we assembled a dextramer pool containing pRBM14, pRAF1, pMYO9B, pZNF518A, pWNK, pSTMN, and pNCOR in complex with C*07:01. In two C*07:01 healthy donor buffy coats we could not detect responses prior to or after sequential dextramer enrichment (figure 5K). These data suggest that although these C*07:01 phosphopeptides are in high abundance in the immunopeptidome,[32] they do not generate prevalent T-cell responses in the autologous setting, and are thus similar to A*03:01 phosphopeptides. A limitation of these studies is that we did not assess these responses in relevant tumor-bearing patients expressing these alleles because we were unable to procure appropriate patient samples. In such patients, the frequencies of PP-CTL may be significantly altered in tumor or peripheral blood presumably due to higher phosphopeptide antigen burden. Another limitation is that we did not survey a large enough number of donors under identical conditions to determine the average relative frequencies of individual phosphopeptide T-cell responses for different alleles. Instead, herein we provide illustrative examples of the best responses seen for each allele from a selected pool and compare them to each other to make conclusions about relative immunogenicity.

DISCUSSION

Phosphopeptides represent an emerging class of HLA ligands that have potential to be targeted as shared tumor antigens produced not by mutations, but by cancer associated post-translational modifications. Since T-cell responses to phosphopeptides have been studied comprehensively in the context of HLA-A*02:01 and HLA-B*07:02,[10 35 46] we sought to expand the phosphopeptidome by isolating phosphopeptides presented by healthy and neoplastic tissue and by systematically reanalyzing large immunopeptidomics data sets containing phosphopeptides, yielding a new set of phosphopeptides presented by the HLA-A3 supertype as well as C*07:01. We further used healthy tissue databases[37] to determine the extent to which these phosphopeptides are found on healthy tissue. Our observation that there were phosphopeptides whose HLA binding adheres to HLA-A3 supertype classifications while being presented exclusively by malignant cells supports the potential of phosphopeptides as tumor antigens that could be targeted across multiple alleles. Supertype binding has been observed previously for phosphopeptides, such as in the case of the same epitope of pIRS2 binding to A*02:01 and A*68:02, and its length variant binding A*03:01 and A*11:01,[47] the latter of which we find in our data as well (table 1). Functionally although aberrant phosphorylation is considered a hallmark of cancer, the phosphopeptides we identified are mostly derived from proteins in common essential pathways, such as nucleic acid binding and repair, rather than oncogenic kinase pathways. In this regard, it is notable that differential presentation of phosphopeptides has been ascribed more to inhibition of critical phosphatases than kinase overactivation.[47]

We used TAP-deficient cells to verify that the MS-identified phosphopeptides that we isolated could stabilize their cognate HLA alleles. All phosphopeptides selected for A*03:01 and C*07:01 bound their cognate allele; however, A*03:01 phosphopeptides did not exhibit a binding advantage compared with their unmodified counterparts. Previously, the half-lives and IC50 values of phosphopeptide-HLA complexes were found to be improved over unmodified counterparts for only 1/5 HLA-A*01:01 peptides, 0/5 HLA-B*07:02 peptides, and 0/7 HLA-B*40:02 peptides.[32 42] This is in contrast to studies of HLA-A2, where phosphorylation was found to increase the binding affinity of a given peptide by 1.1–158.6-fold in 10/11 cases.[9] However, it must also be recognized that the discrepancies in binding properties of pairs of phosphopeptide versus unmodified peptides for their presenting HLA alleles may be affected by technical limitations. Specifically, MS data acquired in DDA mode will contain increased representation of charged peptides, which fragment more efficiently. Since HLA-A2 alleles prefer hydrophobic anchor residues, MS-identified HLA-A2 phosphopeptides may be more likely to contain suboptimal anchor residues that preclude HLA binding in the absence of phosphorylation. This technical limitation may be the underlying cause for identification of phosphopeptide-unmodified pairs that are co-presented by HLA-A3, which prefers charged anchor residues. Interestingly, in phosphopeptide-unmodified

pairs, phosphopeptides more frequently displayed enhanced binding to HLA-C*07:02 and HLA-C*06:02 compared with HLA-A and HLA-B alleles.[32] In our data, we found that most of our selected C*07:01 phosphopeptides were not co-presented with their unmodified counterparts, but the two phosphopeptide-unmodified pairs that were presented did not display any differences in energetics based on molecular docking, suggesting that while co-presentation can occur, the phosphopeptide is the more abundantly presented of the pair. These data highlight the importance of considering the residue preference of the HLA allele of interest when selecting phosphopeptides with augmented binding properties.

Our structural studies based on molecular docking determined that peptide sidechain interactions with the HLA B-pocket and α1 helix were critical determinants of phosphopeptide binding to HLA-A3 and HLA-C7. Previously, HLA-A0201 phosphopeptide binding was shown to be dependent on phosphate-mediated contacts made with Arg65 of the α1 helix.[9] A more recent study found that in HLA-B0702 phosphopeptide complexes, the phosphate moiety was within H-bond distance to Arg62 of the α1 helix.[48] In the case of HLA-A0301, we found that an unmodified P4Ser could mediate more favorable contacts with Asn66 of the α1 helix than phosphoserine. However, attenuated binding was observed in the absence of these interactions in HLA-A3/pMGAP. In a more dramatic case, we compared phosphopeptide binding to HLA-C*07:01 and HLA-C*07:02, alleles which only differ by two B-pocket residues at 66 and 99.[49] We found that Tyr99 of C0701 forms hydrogen bonds with P2 that were absent due to the presence of Ser99 of C0702. These results demonstrated that B-pocket and α1 helix interactions can shape the character of peptide binding to HLA in the presence of subtle differences like phosphate modification or residue substitutions. Our results obtained by molecular docking are qualified, however, by the absence of solvent interactions, which are accounted for in X-ray crystallography and explicit-solvent molecular dynamics.

A feature of phosphopeptides that motivates development of phosphopeptide-targeted agents is that the phosphorylated epitope sequences present a different recognition surface than their unmodified counterparts. Thus, several studies have shown that T cells, TCRs, or TCR-like antibodies specific for MHC-presented phosphopeptides specifically recognize phosphate moieties without cross-reactivity to unmodified peptides.[10 17 40 48] In contrast, class II-restricted pWED-specific T cells did not differentiate between phosphopeptide and unmodified counterpart.[50] Therefore, TCR-like agents capable of phosphorylation discrimination can be developed, but such discrimination is not guaranteed when evaluating native T cells, necessitating the use of adequate methodologic procedures to generate phosphorylation-specific candidates.

Our study demonstrates that several phosphopeptides are selectively presented by several tumor types, but nuances between alleles present important considerations for the development of cancer immunotherapies. Namely, A*11:01 and A*03:01, like B*07:02, are similar in their capacity to co-present phosphopeptides and their unmodified peptides, whereas A*02:01 generally exhibits a binding preference for phosphopeptides. However, phosphopeptides presented by A*11:01, like those of A*02:01 and B*07:02, are immunogenic, whereas phosphopeptides presented by A*03:01 and C*07:01, although demonstrated to be presented, failed to elicit responses in our hands. Alternative approaches for the sensitization of T cells to A*03:01-presented phosphopeptides, synthetic TCRs, or TCR-mimicking antibodies may be required for the generation of PP-CTL or redirection of existing T cells for phosphopeptide-targeted immunotherapy.

Author affiliations
[1]Immunology Program, Memorial Sloan Kettering Cancer Center, New York, New York, USA
[2]Department of Hematology, Oncology and Tumor Immunology, Charite Universitatsmedizin Berlin, Berlin, Germany
[3]German Cancer Research Center, Heidelberg, Baden-Württemberg, Germany
[4]Berlin Institute of Health at Charité –Universitätsmedizin Berlin, BIH Biomedical 13 Innovation Academy, BIH Charité Clinician Scientist Program, Berlin, Germany
[5]Department of Pediatrics, Molecular Pharmacology Program, Memorial Sloan Kettering Cancer Center, New York, New York, USA
[6]Department of Pediatrics, Memorial Sloan Kettering Cancer Center, New York, New York, USA
[7]Weill Cornell Medicine, New York, New York, USA

Acknowledgements We thank the Dr Henrik Molina of the Proteomics Resource Center at Rockefeller University for the performance of all LC/MS-MS experiments. Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the UCSF, with support from NIH P41-GM103311. We acknowledge the use of the MSK Flow Cytometry Core Facility, funded in part through the NIH/NCI Cancer Center Support Grant P30 CA008748. We thank Dr Joshua Elias (Chan Zuckerberg Biohub, Stanford, California, USA) for providing HLA typing of samples from PXD004746 and PXD005704.

Contributors All authors made substantial contributions to the study. ZM conceived and designed the study, developed methodology, acquired data, analyzed and interpreted data, and wrote the original draft of the manuscript. MGK and TD developed methodology, acquired data, and analyzed and interpreted data. JU acquired data. ZM, MGK, TD, JU, DAS, and RJO reviewed and revised the manuscript. DAS and RJO supervised the study. RJO served as guarantor.

Funding ZM and RJO acknowledge support from the Steven A. Greenberg Lymphoma Research Award (GC-242236) and Alex's Lemonade Stand Foundation Innovation Award (GR-000002624). RJO was supported by NIH NCI P01 CA023766, NCI Cancer Center support grant P30 CA008748, Richard "Rick" J. Eisemann Pediatric Research Fund, The Tow Foundation, The Aubrey Fund, and Edith Robertson Foundation. DAS was supported by NIH NCI P01 CA23766 and R35 CA241894. TD was supported by NCI 1R50CA265328. MGK is participant in the BIH Charité Clinician Scientist Program funded by the Charité – Universitätsmedizin Berlin, and the Berlin Institute of Health at Charité (BIH).

Competing interests DAS is on a board of, or has equity in, or income from: Atengen, Lantheus, Sellas, Iovance, Pfizer, Actinium Pharmaceuticals, Inc., OncoPep, Repertoire, Sapience, and Eureka Therapeutics. TD is a consultant for Eureka Therapeutics. MGK is a consultant to Ardigen. RJO declares consultancy, research support, and royalties from Atara Biotherapeutics.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by MSKCC IRB protocol #06-107. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information. Original data sets will be made available upon reasonable request to the corresponding author (Zaki Molvi, Memorial Sloan Kettering Cancer Center, New York, New York, USA; zaki.molvi@gmail.com)

ORCID iDs
Zaki Molvi http://orcid.org/0000-0001-6728-4835
Martin G Klatt http://orcid.org/0000-0001-8703-7305

REFERENCES

1 Hont AB, Cruz CR, Ulrey R, *et al*. Immunotherapy of relapsed and refractory solid tumors with ex vivo expanded multi-tumor associated antigen specific cytotoxic T lymphocytes: a phase I study. *J Clin Oncol* 2019;37:2349–59.
2 Lulla PD, Naik S, Vasileiou S, *et al*. Clinical effects of administering leukemia-specific donor T cells to patients with AML/MDS after allogeneic transplant. *Blood* 2021;137:2585–97.
3 Hays P. Cancer Immunotherapies. In: Hays P, ed. *Allogeneic tumor antigen-specific T cells for broadly applicable adoptive cell therapy of cancer BT - cancer immunotherapies: solid tumors and hematologic malignancies*. Cham: Springer International Publishing, 2022: 131–59.
4 Bassani-Sternberg M, Bräunlein E, Klar R, *et al*. Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass Spectrometry. *Nat Commun* 2016;7:13404.
5 Zarling AL, Ficarro SB, White FM, *et al*. Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo. *J Exp Med* 2000;192:1755–62.
6 Malaker SA, Penny SA, Steadman LG, *et al*. Identification of glycopeptides as posttranslationally modified neoantigens in leukemia. *Cancer Immunol Res* 2017;5:376–84.
7 Zarling AL, Polefrone JM, Evans AM, *et al*. Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. *Proc Natl Acad Sci U S A* 2006;103:14889–94.
8 Depontieu FR, Qian J, Zarling AL, *et al*. Identification of tumor-associated, MHC class II-restricted phosphopeptides as targets for immunotherapy. *Proc Natl Acad Sci U S A* 2009;106:12073–8.
9 Mohammed F, Cobbold M, Zarling AL, *et al*. Phosphorylation-dependent interaction between antigenic peptides and MHC class I: a molecular basis for the presentation of transformed self. *Nat Immunol* 2008;9:1236–43.
10 Cobbold M, De La Peña H, Norris A, *et al*. MHC class I-associated phosphopeptides are the targets of memory-like immunity in leukemia. *Sci Transl Med* 2013;5:1–11.
11 Penny SA, Abelin JG, Malaker SA, *et al*. Tumor infiltrating lymphocytes target HLA-I phosphopeptides derived from cancer signaling in colorectal cancer. *Front Immunol* 2021;12:723566.
12 Cafri G, Yossef R, Pasetto A, *et al*. Memory T cells targeting oncogenic mutations detected in peripheral blood of epithelial cancer patients. *Nat Commun* 2019;10:449.
13 Liu S, Matsuzaki J, Wei L, *et al*. Efficient identification of neoantigen-specific T-cell responses in advanced human ovarian cancer. *J Immunother Cancer* 2019;7:156.
14 Kowarz E, Löscher D, Marschalek R. Optimized sleeping beauty transposons rapidly generate stable transgenic cell lines. *Biotechnol J* 2015;10:647–53.
15 Mátés L, Chuah MKL, Belay E, *et al*. Molecular evolution of a novel hyperactive sleeping beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet* 2009;41:753–61.
16 Klatt MG, Mack KN, Bai Y, *et al*. Solving an MHC allele–specific bias in the reported immunopeptidome. *JCI Insight* 2020;5:e141264.
17 Dao T, Mun SS, Molvi Z, *et al*. "A TCR mimic monoclonal antibody reactive with the "public" phospho-neoantigen PIRS2/HLA-A*02:01 complex". *JCI Insight* 2022;7:e151624.
18 Pettersen EF, Goddard TD, Huang CC, *et al*. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 2004;25:1605–12.
19 Shapovalov MV, Dunbrack RL. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* 2011;19:844–58.
20 Gfeller D, Michielin O, Zoete V. Swisssidechain: a molecular and structural database of non-natural sidechains. *Nucleic Acids Res* 2013;41:D327–32.
21 Raveh B, London N, Schueler-Furman O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins* 2010;78:2029–40.
22 Leaver-Fay A, Tyka M, Lewis SM, *et al*. Rosetta3: an object-oriented software suite for the simulation and design of macromolecules. In: Johnson ML, ed. *Computer methods, part C*. Academic Press, 2011: 545–74.
23 Wölfl M, Greenberg PD. Antigen-specific activation and cytokine-facilitated expansion of naive, human CD8+ T cells. *Nat Protoc* 2014;9:950–66.
24 May RJ, Dao T, Pinilla-Ibarz J, *et al*. Peptide epitopes from the Wilms' tumor 1 oncoprotein stimulate CD4 + and CD8+ T cells that recognize and kill human malignant mesothelioma tumor cells. *Clin Cancer Res* 2007;13:4547–55.
25 Pinilla-Ibarz J, May RJ, Korontsvit T, *et al*. Improved human T-cell responses against synthetic HLA-0201 analog peptides derived from the WT1 Oncoprotein. *Leukemia* 2006;20:2025–33.
26 Dolton G, Tungatt K, Lloyd A, *et al*. More tricks with tetramers: a practical guide to staining T cells with peptide-MHC multimers. *Immunology* 2015;146:11–22.
27 Molvi Z. Peptide-MHC dextramer assembly V1. [Preprint] 2022.
28 Bethune MT, Comin-Anduix B, Hwang Fu Y-H, *et al*. Preparation of peptide–MHC and T-cell receptor dextramers by biotinylated dextran doping. *Biotechniques* 2017;62:123–30.
29 Tsherniak A, Vazquez F, Montgomery PG, *et al*. Defining a cancer dependency map. *Cell* 2017;170:564–76.
30 Narayan R, Olsson N, Wagar LE, *et al*. Acute myeloid leukemia Immunopeptidome reveals HLA presentation of mutated nucleophosmin. *PLoS One* 2019;14:e0219547.
31 Khodadoust MS, Olsson N, Wagar LE, *et al*. Antigen presentation profiling reveals recognition of lymphoma immunoglobulin neoantigens. *Nature* 2017;543:723–7.
32 Solleder M, Guillaume P, Racle J, *et al*. Mass spectrometry based Immunopeptidomics leads to robust predictions of phosphorylated HLA class I ligands. *Mol Cell Proteomics* 2020;19:390–404.
33 Kuleshov MV, Xie Z, London ABK, *et al*. Kea3: improved kinase enrichment analysis via data integration. *Nucleic Acids Res* 2021;49:W304–16.
34 Bern M, Kil YJ, Becker C. Byonic: advanced peptide and protein identification software. *Curr Protoc Bioinformatics* 2012;Chapter 13:13.
35 Lulu AM, Cummings KL, Jeffery ED, *et al*. Characteristics of immune memory and Effector activity to cancer-expressed MHC class I phosphopeptides differ in healthy donors and ovarian cancer patients. *Cancer Immunol Res* 2021;9:1327–41.
36 Sarkizova S, Klaeger S, Le PM, *et al*. A large peptidome dataset improves HLA class I EPITOPE prediction across most of the human population. *Nat Biotechnol* 2020;38:199–209.
37 Marcu A, Bichmann L, Kuchenbecker L, *et al*. HLA ligand atlas: a benign reference of HLA-presented peptides to improve T-cell-based cancer immunotherapy. *J Immunother Cancer* 2021;9:e002071.
38 Vasaikar S, Huang C, Wang X, *et al*. Proteogenomic analysis of human colon cancer reveals new therapeutic opportunities. *Cell* 2019;177:1035–49.
39 McFarland JM, Ho ZV, Kugener G, *et al*. Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. *Nat Commun* 2018;9:4610.
40 Mohammed F, Stones DH, Zarling AL, *et al*. The antigenic identity of human class I MHC phosphopeptides is critically dependent upon phosphorylation status. *Oncotarget* 2017;8:54160–72.
41 Petersen J, Wurzbacher SJ, Williamson NA, *et al*. Phosphorylated self-peptides alter human leukocyte antigen class I-restricted antigen presentation and generate tumor-specific epitopes. *Proc Natl Acad Sci U S A* 2009;106:2776–81.
42 Alpízar A, Marino F, Ramos-Fernández A, *et al*. A molecular basis for the presentation of phosphorylated peptides by HLA-B antigens. *Mol Cell Proteomics* 2017;16:181–93.

43 Tsai J, Taylor R, Chothia C, *et al*. The packing density in proteins: standard radii and volumes. *J Mol Biol* 1999;290:253–66.
44 Di Marco M, Schuster H, Backert L, *et al*. Unveiling the peptide motifs of HLA-C and HLA-G from naturally presented peptides and generation of binding prediction matrices. *J Immunol* 2017;199:2639–51.
45 Hadrup SR, Bakker AH, Shu CJ, *et al*. Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat Methods* 2009;6:520–6.
46 Engelhard VH, Obeng RC, Cummings KL, *et al*. MHC-restricted phosphopeptide antigens: preclinical validation and first-in-humans clinical trial in participants with high-risk melanoma. *J Immunother Cancer* 2020;8:e000262.
47 Mahoney KE, Shabanowitz J, Hunt DF. MHC phosphopeptides: promising targets for immunotherapy of cancer and other chronic diseases. *Mol Cell Proteomics* 2021;20:100112.
48 Patskovsky Y, Natarajan A, Patskovska L, *et al*. Molecular mechanism of phosphopeptide neoantigen immunogenicity. *In Review* [Preprint] 2022.
49 Rasmussen M, Harndahl M, Stryhn A, *et al*. Uncovering the peptide-binding specificities of HLA-C: a general strategy to determine the specificity of any MHC class I molecule. *J Immunol* 2014;193:4790–802.
50 Longino NV, Yang J, Iyer JG, *et al*. Human CD4+ T cells specific for Merkel cell polyomavirus localize to Merkel cell carcinomas and target a required oncogenic domain. *Cancer Immunol Res* 2019;7:1727–39.