Impact of Post-Translational Modification on MHC Peptide Binding and TCR Engagement

Joey J. Kelly¹, Nathaniel Bloodworth², Qianqian Shao¹, Jeffery Shabanowitz¹, Donald Hunt¹, Jens Meiler³,⁴,⁵, and Marcos M. Pires¹*

¹Department of Chemistry
University of Virginia
Charlottesville, VA, United States

²Division of Clinical Pharmacology, Department of Medicine
Vanderbilt University Medical Center, Nashville, TN, United States

²Institute of Drug Discovery, Faculty of Medicine
University of Leipzig, Leipzig, Germany

³Center for Structural Biology
Vanderbilt University, Nashville, TN, United States

⁴Department of Chemistry
Vanderbilt University, Nashville, TN, United States.

* corresponding author: mpires@virginia.edu
Abstract

The human major histocompatibility complex (MHC) plays a crucial role in the presentation of peptidic fragments from proteins; these peptides can be derived from self-proteins or from non-human antigens, such as those produced by viruses or bacteria. To prevent cytotoxicity against healthy cells, thymocytes expressing T cell receptors (TCRs) that recognize self-peptides are removed from circulation in a process called negative selection. However, post-translational modifications (PTMs) are largely excluded from negative selection; this feature opens the door to the possibility that PTMs directly contribute to the development of autoreactive T cells and subsequent autoimmune diseases. Despite it being well-established that PTMs are prevalent in peptides presented on MHCs, the exact mechanisms by which PTMs influence the antigen presentation machinery remains poorly understood. In our work, we introduce chemical modifications mirroring PTMs onto peptides to systematically investigate their impact on MHC binding and TCR recognition. Our findings reveal the numerous ways PTMs alter antigen presentation, which could have implications for tumor neoantigen presentation.
Introduction

To maintain homeostasis, the human immune system must efficiently recognize and destroy cells that have accumulated genetic mutations or have been invaded by pathogenic microorganisms.\(^1\) \(^2\) Self-identification to the immune system is a primary mechanism that is deployed by human cells to flag the presence of non-self-proteins or proteins produced by genetic lesions.\(^3\) To this end, presentation of peptidic fragments from non-self-proteins via the major histocompatibility complex (MHC) to surveying immune cells serves as a key system to recognize diseased cells (Figure 1).\(^4\) MHC is present in the membrane of every nucleated cell and is responsible for presenting both antigenic and endogenous protein fragments to the extracellular space. Binding of peptide-MHC (pMHC) complexes by T cells, which are equipped with T cell receptors (TCRs)\(^5\), initiates an immune cell response; the nature of the response depends on the type of the T cell (primarily CD4+ and CD8+ cells).\(^6\)
Figure 1. Schematic representation of the process involving the proteolytic processing of cytosolic protein into peptides. Subsequent steps lead to the loading of the peptides onto MHC molecules that are then transported for presentation on the cell surface.

Critically, the binding of pMHC by TCRs must operate with a high level of selectivity since the recognition of self-peptides in healthy cells can result in cellular injury. Through negative selection, thymocytes expressing TCRs that bind tightly to self-peptides are removed from circulation. In theory, changes to the primary sequence of a protein could yield autoreactive pMHC as long as that primary sequence was not part of the negative selection process. Some primary sequence changes can be permanent (e.g., amino acid change), while others can be transient, such as post-translational modifications (PTMs). In healthy states, enzymatic PTMs are installed by specific sets of enzymes to modulate protein activity, localization, and interaction of the protein with other cellular components. Because PTMs are covalent modifications, most are stable enough to persist through proteasomal processing; therefore, the resulting modified
peptide fragments can potentially be presented in the context of MHC for immune surveillance.\textsuperscript{10}

Remarkably, human proteins containing PTMs were shown to be excluded from negative selection, despite the potential that pMHC with PTM-modified peptides could yield autoreactive cells.\textsuperscript{11} It was hypothesized that the reason was the relatively low abundance of PTMs in healthy and young adults in the prime phase of negative selection, compared to the unmodified parent proteins. Past adolescence, thymus involution has the potential to contribute to age-related autoimmune diseases.\textsuperscript{12, 13} At the same time, a myriad of factors (balance of PTM writers and erasers in diseased states, cellular stresses, or aging) can dramatically increase the prevalence of PTMs.\textsuperscript{14} Combined, the increased levels of pMHC complexes bearing PTMs means that a higher level of potentially autoreactive cells are present.\textsuperscript{15-21} In other words, peptides with PTM modifications could potentially be interpreted by the immune system as nonself, and, therefore, induce a pathogenic autoimmune response.\textsuperscript{22-24} A conclusive example of this phenomenon was recently described in which cysteine carboxyethylation by a cellular metabolite resulted in a pathogenic neoantigen presented on pMHC that resulted in autoreactive T cell responses.\textsuperscript{25}

A wide range of PTMs have been previously found in pMHCs.\textsuperscript{26-34} As a prominent example, increased conversion of arginine to citrulline in the myelin sheath has been shown to lead to the development of self-reactive T cells that exacerbate the
progression of autoimmune diseases such as Rheumatoid Arthritis (RA) and Multiple Sclerosis (MS). Additionally, there is mounting evidence that PTMs on peptides within pMHC complexes are prevalent in tumors. A recent report revealed that PTMs can shape the antigenic landscape. Despite the importance of PTMs to potentially generate autoreactive peptides, there are many aspects of PTM biology that remain poorly defined in this field. A primary barrier to better understanding the impact of PTMs on pMHC complexes has been the complexity in assigning the role of PTMs in the context of the many steps leading to MHC presentation. These roles can include changes in protein half-live, altered proteolytic processing to presentable peptides, or binding affinity of individual peptides to MHC molecules. Each of these factors on their own has the potential to impact levels and types of peptides displayed on MHCs; of which, binding affinity is a primary factor in defining the repertoire of peptides within pMHCs, given the number of possible peptides that can theoretically be loaded onto MHCs by the transporter associated with antigen processing (TAP). Herein, we sought to systematically isolate the impact of PTMs on the ability of the peptide to bind to the MHC molecules and be recognized by cognate T cells.

Results and Discussion

In general, PTMs can be challenging to detect in the context of complex biological systems. Additionally, the levels of PTMs are dynamic and can shift in response to cellular queues. Traditional cellular assays (e.g., Western blots) do not usually
establish the ratio of PTM proteins relative to their unmodified counterparts. These challenges become amplified when analyzing PTMs on peptides from pMHC complexes. Extraction of peptides from isolated MHCs involves difficult separation steps and downstream mass spectrometry analysis. On the other hand, a top-down approach of treating cells with PTM-modified proteins also poses a challenge because exogenous proteins are not readily processed to yield peptides that are presented by MHC I molecules. Even if the peptides get processed, the ultimate pMHC presentation entails a combination of factors including peptide half-life, uptake efficiency, processing by proteases, or peptide binding affinity to MHC molecules. Therefore, identifying how various naturally occurring PTMs impact peptide binding to MHCs and recognition by TCRs has not been carried out in a systematic way.

To overcome these challenges, we centered our approach on the RMA-S cell line that lacks the TAP importer. TAP is essential for delivering peptides from the cytosol to the endoplasmic reticulum. Once in the endoplasmic reticulum, high affinity binding of peptides to MHC leads to stable associations of this complex and robust presentation on the cell surface. In contrast, TAP deficient cells have inherently low levels of MHC class I expression on the cell surface. Lowering the temperature (22-26 °C) allows for proper expression of MHC class I that are loaded with low affinity peptides or are empty (Figure S1). Cell treatment with peptides that bind MHC with high affinity results in pMHC complexes with thermostability and high surface presentation. Effectively, peptides in the extracellular space displace low binding peptides or associate with empty MHC molecules. Upon the elevation of the temperature to 37 °C,
Low affinity binders dissociate from the pMHC complexes leading to internalization and degradation of the associated MHC molecules. In contrast, supplementing a high-affinity MHC binding peptide to the culture medium at 26 °C enables the stabilization of pMHC on the surface of the cell when the temperature is raised to 37 °C.

The quantity of MHC on the cell surface correlates with the binding affinity of the peptide to the MHC molecule, which can be quantified via detection with a fluorescently labeled anti-MHC antibody. Given these features of this well-established cell line, we anticipated that the RMA-S cells would serve the role of revealing the impact of PTMs on peptides that bind to MHC with high affinity. Moreover, there is considerable precedence for using RMA-S cell lines for isolating the affinity of a peptide for MHC class I in contexts other than PTMs. To start, we sought to identify a peptide from SARS-CoV-2 spike protein. We used NetMHCPan4.0, an algorithm that estimates the propensity of peptides to bind MHC, to select a peptide sequence (ESIVRFPNI, referred to as sarsWT) for stabilization of the specific MHC molecule (H-2Kb) on the surface of RMA-S cells.

The peptide was synthesized using standard solid-phase peptide synthesis techniques. The culture medium of the RMA-S cells was supplemented with sarsWT (Figure S2A) at 26 °C to enable association of the peptide with surface bound MHCs. Following sarsWT incubation, the temperature was raised to 37 °C. Subsequently, the cells were incubated with APC conjugated anti-mouse H-2Kb antibody and analyzed by
flow cytometry where fluorescence is expected to correspond with the amount of sarsWT bound to the MHC. The peptide SNFVSAGI (cntPEP) was used as a negative control since it has been reported to not interact appreciably with H-2Kb. Various concentrations of sarsWT were used to optimize the RMA-S stabilization assay with the goal of increasing the number of pMHC molecules formed on the surface of the cells (Figure S2). We found an increased fluorescence signal corresponding to increasing concentrations of sarsWT, indicating that the peptide bound and stabilized the MHC on the surface of the cell. The amount of stable pMHC complexes that formed at the surface also saturated at 50 μM of peptide and showed an EC50 of 1.7 μM. These results confirmed the suitability of RMA-S to report on the stabilization of the pMHC complex by sarsWT.

Figure 2. (A) Chemical structures of sarsWT and the PTM modified variants. (B) Flow cytometry analysis of RMA-S cells treated with specific peptide (20 μM) detected by
APC conjugated anti-mouse H-2K\textsuperscript{b} antibody. Data are represented as mean ± SD (n=3). P-values were determined by a two-tailed \textit{t}-test (* \( p < 0.05 \), ** \( p < 0.01 \), *** \( p < 0.001 \), **** \( p < 0.0001 \), ns = not significant).

Another parameter to consider is the time in which the MHC binding peptide has to associate with the MHC to reach MHC saturation. To evaluate this time frame, the workflow described above was conducted; however, the peptide incubation time at 37 °C was varied from 2 to 6 h (Figure S3). The results showed that the fluorescent signal increased with increased incubation periods. This is likely due to accumulation of stabilized pMHC complexes in the presence of excess high-affinity peptide. Nevertheless, a 6 h incubation at 37 °C provided the highest signal-to-noise ratio and was used for downstream assays.

With an optimized protocol in hand, we then synthesized a panel of peptides containing PTMs to investigate their potential impact on MHC binding (Figure 2A). We envisioned that modifying residues on \textit{sarsWT} would be representative of how the selected modifications may impact binding of different MHC peptides. The specific PTMs chosen were a proline to hydroxyproline\textsuperscript{61, 62}, arginine to citrulline\textsuperscript{63}, and N-terminus acetylation\textsuperscript{64} modifications, as they are naturally occurring PTMs. By performing the RMA-S stabilization assay, we found that the citrullination and hydroxyproline modifications had no significant effect on MHC binding, but the N-terminal acetylation modification completely disrupted MHC binding (Figure 2B). Gratifyingly, there is precedence for N-acetylated peptides to be displaced in MHC class I.\textsuperscript{65, 66}
These results demonstrate that naturally occurring PTMs can alter pMHC complex formation and influence the antigen presentation pathway depending on the interface of the peptide and the MHC molecule (Figure 3).

Figure 3. Schematic representation of how the orientation of the PTM within the presented peptide could negatively impact binding to MHC molecules (right) compared to the unmodified peptide (left).

We then shifted our focus to the peptide epitope SIINFEKL (ovaWT) from the model antigen ovalbumin (OVA) to interrogate how other PTMs impact MHC binding as well as TCR recognition. To this end, we synthesized a new library of OVA peptides containing PTMs on its serine and lysine residues and performed a concentration scan to assess MHC binding using the RMA-S stabilization assay (Figure 4A). Interestingly,
there was a significant decrease in MHC binding when serine was phosphorylated despite predictions that this site does not contribute significantly to binding. Our results highlight the potential deficiency of current models to predict how PTMs could impact peptide binding to MHC molecules given how PTM products can be structurally distinct relative to the sidechains of the canonical amino acids.

Figure 4. (A) Chemical structures of ovaWT and the PTM modified variants. (B) Flow cytometry analysis of RMA-S cells treated with specific peptide (20 μM) detected by APC conjugated anti-mouse H-2Kb antibody. Data are represented as mean ± SD (n=3). P-values were determined by a two-tailed t-test (* p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001, ns = not significant).

The sidechain of lysine on ovaWT was explored next. Lysine was decorated with different types of PTMs that included three different charge states. Previously, the same
lysine on SIINFEKL had been altered with various groups including fluorophores and caging groups. Our data showed considerable accommodation of structural alterations to the lysine sidechain (Figure 4B). With larger modifications such as biotin, binding levels are altered but the change is not considerable. These findings can be explained by the solvent exposed nature of lysine 7 on the pMHC complex and the length of the lysine sidechain, which may reduce structurally unfavorable interactions with the MHC molecule (Figure 3). Critically, for ovaWT the anchor residues that are more important for the overall binding affinity to MHC molecules (Phe5 and Leu8) are not amendable to conventional PTMs. While PTMs on solvent exposed residues could potentially retain binding to MHCs, their overall impact may be more strongly reflected in the pMHC engagement with TCRs. TCRs recognize T cell exposed motifs of the bound peptide within pMHCs. Therefore, PTMs at these positions may have a greater impact on T cell activation.

To assess the impact of ovaWT PTMs on TCR recognition, we utilized the B3Z T cell hybridoma cell line that contains an OVA specific TCR along with a NFAT-LacZ reporter gene that encodes for β-galactosidase on an IL-2 inducible promoter. Upon B3Z recognition of the OVA pMHC complex on RMA-S cells, IL-2 production promotes β-galactosidase expression, which can hydrolyze chlorophenol red-β-D-galactopyranoside (CPRG), and the resulting color change is representative of activation levels. Our results showed that the PTMs have a much more significant effect on TCR recognition than on peptide binding to MHC (Figure 5A). While relatively modest disruption of TCR recognition was seen for increasing the degree of methylation on the lysine residue,
near complete disruption of TCR recognition was seen for all PTMs that imparted a change in charge of either the serine or lysine residue. Additionally, the reduction in TCR recognition for both the phosphoserine and biotinylated OVA cannot be fully explained solely by their decrease in MHC binding. Instead, the PTMs that are well accommodated for MHC binding could be displayed away from the binding cleft and alter engagement with TCRs on cognate T cells (Figure 5B).

Figure 5. (A) RMA-S cells were incubated with peptide and B3Z T cells overnight at an effector to target ratio of 1:1. β-galactosidase expression was then measured via the colorimetric reagent CPRG on a plate reader at 570 nm.

To confirm the relevance of these findings in a broader context, we also performed the T cell activation assay on DC2.4 dendritic cells, which also express H-2Kb. Unlike the RMA-S cells, peptides were incubated with DC2.4 cells at 37 °C and were expected to
load onto MHC molecules if they display sufficient affinity towards H-2Kb. Satisfyingly, the pattern of T cells activation with PTM modified ovaWT peptides showed a similar profile with DC2.4 cells as RMA-S cells (Figure S4). Our results confirmed that the impact of PTMs on T cell activation may be consistent across multiple cell types. Finally, we sought to investigate whether we could assess PTMs on a peptide that has potential pathological implications. To this end, a number of autoimmune diseases, including RA and MS, have been described to involve the citrullination of arginine. Citrullination is carried out by peptidylarginine deiminases (PADs) and higher levels of citrullinated proteins have been found in older mice relative to young mice. A shift in higher levels of citrullination past the full scope of negative selection could provide a pathway for autoreactivity. Using a peptide originating from a known PAD substrate myelin basic protein (RTAHYGSL, mbpWT) as the baseline (Figure 6A), we found that there was a statistically significant increase in peptide binding to MHC molecules upon citrullination. This result is consistent with the possibility that PTMs can impact presentation of peptides in the context of disease-linked peptides.
Figure 6. (A) Chemical structures of mbpWT and the PTM modified variant. (B) Flow cytometry analysis of RMA-S cells treated with specific peptide (20 μM) detected by APC conjugated anti-mouse H-2Kb antibody. Data are represented as mean ± SD (n=3). P-values were determined by a two-tailed t-test (* p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001, ns = not significant).

To gain additional insight into how PTMs might alter peptide binding affinity to H-2Kb, we simulated peptide binding to H-2Kb using ROSETTA and FlexPepDock. FlexPepDock has been previously benchmarked against MHC-I bound peptides and is capable of generating models with sub-angstrom accuracy.\textsuperscript{79, 80} One advantage of FlexPepDock over contemporary machine learning methods is its ability to incorporate non-canonical amino acids and PTM residues by generating custom ROSETTA parameters files; this enables FlexPepDock to generate accurate models of MHC-I bound epitopes containing such residues.\textsuperscript{81, 82} Furthermore, the binding energy metrics...
calculated by this application (the reweighted_sc score term) can serve a surrogate for peptide binding affinity to MHC-I.\textsuperscript{83} We simulated the binding of the SARS-CoV-2-derived peptide to H-2K\textsuperscript{b} and generated results recapitulating experimentally derived binding affinity changes (\textbf{Figure 7}). ROSETTA correctly predicts decreased binding affinity for the top-scoring decoys after N-terminal acetylation (corresponding to an increase in the average reweighted_sc term; ROSETTA energy metrics are inversely proportional to binding affinity). ROSETTA predicts either no change or a modest increase in binding affinity for the other two PTMs assessed (\textbf{Figure 7A}). When the top-scoring model for each peptide is visually inspected, we noted significant alterations in the peptide backbone structure for the acetylated N-terminus peptide variant largely confined to the H-2K\textsuperscript{b} N-terminus binding pocket (\textbf{Figure 7B}), a region that is largely responsible for dictating epitope binding specificity and that cannot easily accommodate large conformational changes.\textsuperscript{9} Conversely, R5 citrullinated and P7 hydroxylated variants are located outside the H-2K\textsuperscript{b} binding pockets and regions tolerant of larger conformational changes (\textbf{Figure 7C}).
Figure 7. Modeling the sars peptide with and without PTMs using FlexPepDock. (A) N-terminal acetylation results in higher reweighted_sc values (and thus lower likelihood of binding) compared to the wild-type peptide, consistent with experimental results (one-way ANOVA and Holm-Sidak post-hoc; individual values correspond to the top 0.5% of models generated, and bars represent mean score term values). (B) Top peptide models generated for each epitope positioned in the H-2Kb binding cleft. (C) Side chain modifications compared to the unmodified epitope for the SARS-CoV-2 peptide.
We repeated this analysis on the ovalbumin-derived peptide and each PTM (Figure 8A). Phosphorylation of the N-terminus serine generates structures with higher (and therefore less favorable) score terms, again likely due to disruption of the N-terminus peptide-binding pocket (Figure 8B). Peptide models with the remaining PTMs scored either equivalently or slightly better than their wild-type counterparts. While largely consistent with experimental data, ROSETTA predicts modest increases in peptide binding favorability that are not observed experimentally. This may be a result of biases inherent in ROSETTA’s scoring weights that have not been fully optimized for unusual or non-canonical amino acids such as those modeled here. Alternatively, the assay used for assessing binding affinity may be insufficiently sensitive to detect small changes above certain affinity thresholds. Modifications at the lysine in position 7 alter T cell activation in vitro as illustrated in Figure 5. In silico, these altered residues do not interact appreciably with peptide binding pockets but instead occur in the region displayed by MHC-I to patrolling T cells (Figure 8C). Methylated and acetylated lysine alters residue charge and hydrophobicity; succinylation adds a negative charge (as opposed to lysine’s normally positively charged side chain); and biotinylation results in the addition of a large aliphatic heteropolycyclic side chain very different in size and character from the native lysine residue.
Figure 8. Modeling the ovalbumin-derived peptide with and without PTMs using FlexPepDock. (A) Phosphorylation of the N-terminal serine residue generates models with higher average reweighted_sc terms, similar to N-terminal acetylation of the SARS-CoV-2 peptide (one-way ANOVA and Holm-Sidak post-hoc; individual values correspond to the top 0.5% of models generated, and bars represent mean score term values). (B) Superimposed peptide backbones for the top peptide models generated for
each PTM. (C) Detailed side-chain configurations for the top-scoring models generated by FlexPepDock.

Finally, we performed a similar analysis with the mbp peptide with and without R1 citrullination and generated in silico results that recapitulated experimental findings (Figure 6). Citrullination at R1 significantly increases the reweighted_sc metrics for the models generated by ROSETTA (Figure S5A). While this modification occurs at the N-terminus binding pocket, it does not appreciably alter peptide backbone configuration (Figure S5A, in contrast to the effects of N-terminal acetylation of the sars peptide). It does, however, alter side chain hydrophobicity and charge. Given that the citrullinated variant is charge-neutral (compared to arginine’s 1+ charge) and given H-2Kb’s preference for hydrophobic or charge-neutral residues at binding pockets, it is reasonable to expect this modification may enhance peptide binding affinity.

Increasingly, there are efforts dedicated to developing therapeutic agents that harness a patient’s immune system against cancerous lesions. A prominent example involves the modulation of the programmed death-1 (PD-1) system and its cognate programmed death-ligand 1 (PD-L1). Cancers can often hijack this set of proteins to maintain growth while suppressing the patient immune response against transformed cells. Immune checkpoint inhibitors that disrupt the association of PD-1 and PD-L1 have shown potent anti-cancer activity by a number of mechanisms but primarily via increased T cell engagement. Neoantigens arising from genetic lesions likely provide a pool of antigens that can be presented by cancer MHCs on the cell surface during checkpoint
therapy. Importantly, our results suggest that therapeutic agents that increase the presentation of peptides through an altered balance of PTM marks could potentially complement the targeting antigenic pool.

Considering the observed roles of PTMs in MHC binding and recognition by TCRs, we propose that the balance of PTM addition by “writers” and PTM removal by “erasers” can be central in immunological health and in diseased states. This imbalance could be triggered by age or manufactured by therapeutic interventions. For example, proteins could naturally exist in high abundance of non-acetylated states on lysine side chains past the peak period of negative selection of T cells. After cancer development, there is an increased level of acetylation, but those marks need to be removed by “erasers” called Lysine Deacetylases (KDAC) to reduce an anticancer immunological response against the cancerous cells. Upon the administration of KDAC inhibitors, a larger pool of acetylated peptide would be presented. As our data showed, the acetylation of lysine that interacts with TCRs would escape negative selection and provide an anticancer immunological response. Such ‘cryptic’ antigens could potentially be driving the pharmacological effect of some PTM modulators in cancer patients. To this end, there are a high number of ongoing studies evaluating the combination of KDAC inhibitors and PD-1/PDL-1 inhibitors across many types of difficult-to-treat tumors.\(^8^9-^9^5\) We are currently working towards using our platform to demonstrate how KDAC inhibitors directly lead to an immunopeptidome that reveal cryptic antigens able to enhance anti-tumor immunological activity.
Conclusion

In conclusion, the data presented here demonstrate that naturally occurring PTMs can drastically impact the antigen presentation pathway by altering either affinity towards MHC molecules or TCR recognition by cognate T cells. In our assays, we showed that some PTMs can disrupt peptide binding to MHC, but primarily on residues that directly engage with MHCs. Conversely, PTMs on residues whose sidechains are positioned away from the MHC binding face are more sensitive to TCR recognition due to the altered binding modality with TCRs. To the best of our knowledge, our study is the first systematic analysis of the impact that PTMs have on MHC binding in a whole cell context and how PTMs can also alter T cell activation using live target and effector cells. Provided that negative selection is age-dependent and becomes increasingly diminished past adolescence, it is evident that PTMs could be causal events that result in autoimmune diseases.

References


