Published in final edited form as:

J Proteome Res. 2017 January 06; 16(1): 228–237. doi:10.1021/acs.jproteome.6b00496.

Identification and Characterization of Complex Glycosylated Peptides Presented by the MHC Class II Processing Pathway in Melanoma

Stacy A. Malaker^{1,6}, Michael J. Ferracane⁵, Florence R. Depontieu⁴, Angela L. Zarling^{3,7}, Jeffrey Shabanowitz¹, Dina L. Bai¹, Suzanne L. Topalian⁴, Victor H. Engelhard³, and Donald F. Hunt^{1,2,†}

¹Department of Chemistry, University of Virginia, Charlottesville, VA 22901

²Department of Pathology, University of Virginia, Charlottesville, VA 22901

³Beirne B. Carter Immunology Center Department of Microbiology, University of Virginia, Charlottesville, VA 22901

⁴Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD 21205

⁵Department of Medicinal Chemistry, University of Florida, Gainesville, FL 32610

Abstract

The MHC class II processing pathway presents peptides derived from exogenous or membrane-bound proteins to CD4⁺ T cells. Several studies have shown that glycopeptides are necessary to modulate CD4⁺ T cell recognition, though glycopeptide structures in these cases are generally unknown. Here we present a total of 93 glycopeptides from three melanoma cell lines and one matched EBV-transformed line, with most found only in the melanoma cell lines. The glycosylation we detected was diverse and comprised 17 different glycoforms. We then used molecular modeling to demonstrate that complex glycopeptides are capable of binding the MHC and may interact with complementarity determining regions. Finally, we present the first evidence of disulfide-bonded peptides presented by MHCII. This is the first large scale study to sequence glyco- and disulfide bonded MHCII peptides from the surface of cancer cells, and could represent a novel avenue of tumor activation and/or immunoevasion.

TOC image

[†]to whom correspondence should be addressed. Telephone: (434)-924-3610 dfh@virginia.edu.

⁶Current address: Department of Chemistry, Stanford University, Stanford, CA 94035

⁷Current address: PPD Laboratories, Richmond, VA 23220.

Supporting information: The following files are available free of charge at ACS website: <http://pubs.acs.org>:

Keywords

immunology; mass spectrometry; glycopeptide analysis; MHC class II

Introduction

The major histocompatibility complex (MHC) class II processing pathway presents peptides derived from exogenous or membrane-bound proteins to CD4⁺ T cells. These antigens are internalized via endo- or pinocytosis and are processed by resident proteases to create 10–25 amino acid peptides that are loaded onto MHC class II molecules. The resulting peptide-MHC complex is subsequently exported to the cell surface, where it can interact with T cell receptors (TCRs) of CD4⁺ T cells and initiate their effector and regulatory functions. Activated CD4⁺ T helper cells will also interact with B cells to stimulate the generation of IgG antibodies.^{1, 2}

Post-translationally modified (PTM) peptides constitute a small percentage of MHC class II bound peptides, and a number of modified peptides have been described *in vivo*.^{3, 4} One of the most common PTMs is protein glycosylation, and evidence suggests that glycopeptides can be presented by the class II processing pathway, as several studies have shown that glycopeptides can modulate CD4⁺ T cell recognition. The first example of a carbohydrate-dependent T cell response was provided by Ishioka *et al*, who reported producing glycopeptide-specific T cells in a murine system by conjugating a carbohydrate to a peptide that binds class II MHC molecules.⁵ Shortly following this study, Michaelsson *et al* showed that carbohydrates present on collagen type II (260–280) are essential for its T cell reactivity in rats.⁶ Since then, extensive studies have demonstrated that T cell responses toward this glycosylated antigen are associated with mouse and human rheumatoid arthritis.^{7–9} Finally, Topalian and colleagues provided evidence that glycosylation was necessary for T cell recognition of a known melanoma antigen derived from tyrosinase in human cells.¹⁰

Despite these examples, very few glycopeptides naturally presented by MHC class II have ever been sequenced. In 1993, Chicz *et al* were the first to describe glycopeptides naturally presented by class II human leukocyte antigen (HLA), identifying and sequencing two peptides from LAM Blast 1 modified by *N*-acetylglucosamine (GlcNAc) on Asn104.¹¹ The second set of HLA-associated glycopeptides were described by Dengjel *et al*, who found

two CD3-derived peptides modified by GlcNAc₂-Fucose₁-Mannose₃.¹² This was the first example of a complex glycan being presented by the MHC class II pathway. To our knowledge, these four peptides are the only naturally processed class II glycosylated peptides that have been directly sequenced to date.

Protein glycosylation may affect epitope recognition by T cells in a number of ways. First, glycopeptides may affect enzymatic cleavage of glycoproteins for antigen processing and/or T cell recognition, as this mechanism is not well understood. Second, the glycosylated residue could affect peptide-MHC binding directly or through conformational influence on the peptide-bound moiety.¹³ Finally, the sugar moiety could make direct contacts with the TCR and influence its activity. However, due to the shortage of naturally presented glycopeptides described in the literature, much speculation exists on this potentially important interaction.

Additionally, tumor cells have altered glucose metabolism, as cancer cells preferentially utilize glycolysis to produce ATP in order to satisfy their increased energetic and biosynthetic requirements. Although aerobic glycolysis is an inefficient process compared to oxidative phosphorylation, in a nutrient rich environment, efficient energy production is less important than biosynthesis. And since cancerous cells are programmed to rapidly proliferate, they must consume carbon and nitrogen-rich nutrients to biosynthesize metabolites needed for cell division. This metabolic shift, termed the “Warburg effect,” has been confirmed in human cancers *in vivo*, as these tumors require a much higher glycolytic flux compared to adjacent normal tissue.¹⁴ Interestingly, the rate of glucose metabolism directly correlates with tumor aggressiveness.¹⁵ As such, we hypothesized that the increased glucose content within cells would result in tumor-specific glycopeptides presented by the MHC class II processing pathway and that these glycopeptides could be attractive antigens for cancer immunotherapies.

Here we describe 93 peptides modified with *O*- and *N*-linked glycosylation that are isolated from three human melanoma cell lines (2048mel, 1102mel, and 1363mel) and one autologous non-melanoma EBV transformed cell line (2048EBV). We subsequently model three of these glycopeptides in complex with HLA class II molecules and TCRs in order to understand how these complex glycans might affect CD4+ T cell recognition. Finally, we identify 30 glycopeptides that also contain a disulfide bonded peptide. Ultimately, we believe the described methods and their accompanying results could be important for understanding the importance of carbohydrate presentation in immune responses against cancer as well as microbial antigens and autoantigens.

Results and Discussion

Identification and Characterization of HLA-DR-Associated Glycopeptides

To identify naturally processed tumor-associated MHC II-restricted glycopeptides as potential targets for immune recognition, we affinity-isolated HLA-DR peptide complexes from three cultured melanoma lines (1102mel, 1363mel, and 2048mel) and one autologous EBV-B cell counterpart (2048EBV). We have previously used these samples for identification of tumor-associated phosphopeptides,¹⁶ and decided to reexamine them due to

our success in the former study. These cell lines are also useful because each cell line – 2048mel (HLA-DR β 1*0101, -DR β 1*0404, and -DR β 4*0103), 1102mel (HLA-DR β 1*0401 and -DR β 1*1502), and 1363mel (HLA-DR β 1*0101) – constitutively expresses substantial levels of common HLA-DR molecules. Importantly, HLA-DR β 1*0101 and HLA-DR β 1*0401 are among the most commonly expressed DR alleles in the melanoma patient population.

Using a higher energy collision activated dissociation (HCD) triggered electron transfer dissociation (ETD) method (see Materials and Methods) on a Thermo Orbitrap Fusion, we identified a total of 93 glycopeptides from the three cell lines (Table 1). HCD relies on collisions between the glycopeptide and nitrogen gas to trigger fragmentation, which causes sequential losses of monosaccharide units from the peptide. Thus, HCD spectra allow for the assignment of glycosylation patterns due to characteristic losses of masses corresponding to individual sugars from the precursor ions. ETD fragments peptides by transferring an electron from a radical anion to a protonated peptide and causes random fragmentation of the peptide backbone. As this fragmentation technique does not rely on collision for fragmentation, PTMs are preserved and site localization of modified residues can be confidently assigned.¹⁷ Thus, the ETD spectra allowed for unambiguous assignment of modified residues and peptide sequence.

In addition to the 93 peptides isolated from human proteins, we detected 11 glycopeptides originating from bovine proteins (SI Table 1), presumably from the fetal bovine serum used in cell culture medium. We have excluded these from the following calculations, but their presence is further evidence that exogenous glycosylated peptide antigens can be presented by the class II pathway. Also, we would like to note that here we report only those peptides that we could unambiguously sequence. Hundreds of other HCD spectra contained the canonical GlcNAc fingerprint and neutral losses associated with sugars, but did not contain enough charge to fragment efficiently by ETD. Thus, the peptides presented here are a small sampling of a much larger glycopeptide population.

Thirteen of the glycopeptides were found in two or more cell lines, yielding a total of 80 unique glycopeptides. Similar to unmodified MHC class II peptides, the average length of these peptides was 17 amino acids (range 11–25). Also typical of MHC class II epitopes, 78% (62/80) were found within nested sets, which are defined as sequences containing the same core sequence but differing in N- and/or C- termini.¹⁸ As such, a total of 26 unique epitopes were detected. Most of the glycopeptides were found only in the melanoma cell lines (91%, 73/80). A small fraction of the glycopeptides (9%, 7/80) was only expressed in EBV transformed cells. One peptide was found in both 2048mel and 2048EBV-B, though the relative abundance was 30 times higher in the matched melanoma cell line. Thus, we have identified 1 tumor-associated and 72 tumor-specific glycopeptides.

The 80 glycopeptides originate from 24 different source proteins. In terms of subcellular location, according to Uniprot, 46% (11/24) are associated with the cell membrane, and 13% (3/24) are associated with the lysosome membrane, and 33% (8/24) are secreted; though most are not truly secreted and are instead associated with the extracellular matrix. Of note, many of these proteins are important for essential cell functions including signal

transduction, receptor signaling, and most commonly, cell adhesion. As such, many are known to be implicated in cancer. For instance, L-dopachrome tautomerase, also known as TRP2, is a well-known melanocyte differentiation antigen and target for melanoma immunotherapy.¹⁹ Additionally, vascular endothelial growth factor receptor 1 (VEGF1) is a receptor tyrosine kinase that acts as a cell-surface receptor for VEGF and plays an essential role in the regulation of angiogenesis, cell survival, cell migration, macrophage function, and cancer cell invasion. A recent review highlights its utility in cancer immunotherapy.²⁰ Finally, intercellular adhesion molecule 1 (ICAM1) expression was shown to have a significant role in mediating cancer cell invasion in lung cancers.²¹ Given the known importance of these proteins, we believe these MHC-associated glycopeptides may prove to be important in the understanding and design of cancer immunotherapies.

The glycosylation we discovered was extremely diverse and occurred on 28 unique sites, most of which have been reported in Uniprot (79%, 22/28). A majority of the glycosylation (82%, 23/28) was *N*-linked (on Asn), whereas a small proportion (18%, 5/28) was *O*-linked (on Thr or Ser). This distinction could be made based on peptide sequence information provided by the HCD and/or ETD spectra.

Canonical *O*-linked glycosylation has been shown to consist of (a) a single *O*-GlcNAc or (b) multiple saccharide units beginning with *O*-GalNAc.²² *O*-GlcNAc is a ubiquitous modification that modulates cellular functions through extensive cross-talk with signaling cascades that are also regulated by phosphorylation. Aberrant *O*-GlcNAcylation correlates with augmented cancer cell proliferation, survival, invasion and metastasis.²³ On the other hand, *O*-GalNAc is normally modified with other carbohydrates such as galactose, GlcNAc, and sialic acid. These large glycans are responsible for the formation of the extracellular matrix and/or mucosal secretions. When the glycans are shortened to *O*-GalNAc or *O*-GalNAc-galactose, they are known as tumor associated carbohydrate antigens (TACAs), which are highly implicated in cancer progression.²⁴ Interestingly, we discovered both *O*-GlcNAc and *O*-GalNAc (Tn antigen) on the HLA class II associated peptides (Table 1 and Figure 1).

Unlike *O*-linked glycosylation, *N*-linked glycosylation occurs in three steps: (1) synthesis of the dolichol-linked precursor oligosaccharide, (2) transfer of the precursor oligosaccharide to the protein, and (3) processing and diversification of the oligosaccharide. Thus, three main types of glycans exist, representing the various stages of maturation. The first, “high mannose”, consists of two GlcNAcs, up to nine mannose residues, and one or more of the initial glucose residues that were part of the lipid-linked oligosaccharide. The second, “complex” oligosaccharides, contain the trimannosyl chitobiose core (GlcNAc₂Mannose₃) along with several other monosaccharide units attached to this core. Finally, “hybrid” oligosaccharides, as implied by their name, are a combination of high mannose and complex sugars.²² Typically, initial processing of the glycan occurs in the endoplasmic reticulum (ER) and are then processed through the Golgi, where the sugars mature. As such, the complex type of glycan is generally associated with proteins that move to the cell surface and may be secreted. We have identified peptides modified with all three types of glycans, truncated or intact, presented by the MHC class II processing pathway (Table 1 and Figure

1). All of the *N*-linked glycosylation sites we identified follow the recognized consensus sequence NX(S/T).

The most common glycosylation that we detected was GlcNAc₂-Fucose₁-Mannose₃, which is the base of all three types of complex glycans. We are currently unable to distinguish whether this structure is processed within the lysosomal compartment and presented as such, or if the truncation occurs during sample handling. However, previous studies have shown that MHC molecule glycosylation remains unchanged during sample processing.¹¹ Thus, these glycans might be truncated in the lysosomal compartment before loading onto the HLA class II molecules and presented to CD4+ T cells. Another possibility is that the glycoproteins are targeted to the lysosome straight out of the ER without going through the trans-Golgi so that they retain their high-mannose structures. Future efforts will be devoted to determining if one or both of these hypotheses is correct.

Analyzing the abundance of the isolated glycopeptides revealed a large range in copy numbers per cell (from 1 to >400), with the average being 37 copies per cell. Interestingly, in our previous study we often detected phosphopeptides at less than 6 copies per cell.¹⁶ This suggests that the individual glycopeptides are actually more abundant than phosphopeptides, although technical differences in isolating and sequencing glycopeptides versus phosphopeptides might also explain this observation. Since T cell responses may be activated by fewer than 10 peptide-MHC complexes per cell,²⁵ the tumor-specific glycopeptides described here could potentially be utilized in melanoma vaccines designed to stimulate anti-tumor immunity.

Modeling of Glycopeptides with HLA class II and TCR

Several possibilities exist as to how the saccharides might affect HLA and/or TCR binding. To begin, we utilized Immune Epitope Database and Analysis Resource (IEDB; <http://tools.iedb.org/mhcii/>) to first predict which nine residues of each of the peptides would likely reside within the MHC class II binding pocket. The allele was assumed to be HLA-DRβ1*0101 if presented by 1363mel, since this is the sole DR allele shown to be expressed by this melanoma cell line. If uniquely expressed by another cell line, the lowest scoring allele was used. The lowest scoring nine residue motif for each peptide, where a low score corresponds to a good binder, is highlighted in bold in Table 1. Those with no score listed were too short to receive an accurate prediction of binding. Of the 26 distinct nine-residue sequences, 23 did not contain the modified residue. We predict that the glycan, in these cases, will not affect MHC-peptide binding. However, 3 peptides are predicted to contain the glycosylated amino acid within the peptide's MHC binding core, as indicated by the asterisk in the scoring column.

We then used molecular modeling to assess how these complex glycans might affect HLA and/or TCR binding. Briefly, crystal structures of five ternary HLA-peptide-TCR complexes^{26–30} and glycans of gp120³¹ were used to construct model HLA-glycopeptide-TCR complexes. The backbone conformation of the α and β HLA subunits (RMSD C_α between 0.45–1.04 Å and 0.54–1.04 Å respectively) and the presented peptide antigens (RMSD C_α between 0.48–2.00 Å) are essentially identical among the crystal structures, allowing the backbone conformation and subsequent placement of the three glycopeptides to

be modeled with high confidence. Glycans, on the other hand, have high conformational mobility, a property that limits their resolution in X-ray crystallography and complicates modeling of their structures.³² Here, the simplified glycan fragments (GlcNAc)₂(Mannose)₃ and (GlcNAc)₂(Fucose)(Mannose)₃ were appended to the constructed peptides (Table 2), and ensembles of conformations were generated to approximate the steric effect imposed by each peptide's glycan(s).

The glycans of the three modeled glycopeptides are attached to residues that flank the HLA peptide binding groove (Figure 2). Though the glycans – as a result of their size and potential effect on peptide conformation – will likely slow initial formation of the HLA-peptide complex, they do not appear to prevent binding of glycopeptides to the HLA. Accordingly, we were able to identify the presence of HLA-bound glycopeptides experimentally, and our modeling suggests that these glycopeptides bind with their hydrophilic glycans projected away from the HLA and into the aqueous extracellular milieu. This positions the modeled glycans, which are often truncated versions of those actually observed experimentally, either into or adjacent to the space occupied by TCRs in a ternary HLA-glycopeptide-TCR complex (Figure 2). This suggests that the large, flexible glycans are capable of adopting conformations in which they can interact (specifically or nonspecifically) with the TCR to prevent or enhance (a) formation of the ternary complex and (b) dissociation of the TCR from the ternary complex. Of course, the effect of these interactions will vary depending upon the identity and location of the glycan as well as the unique sequence and structure of the CDR of each individual TCR. Still, it would be expected that the modeled glycopeptides would possess different activity than their non-glycosylated analogs.

Recent structural work has shown that the super-antigen staphylococcal enterotoxin B (SEB) can bind the HLA-peptide-TCR complex.²⁸ In the crystal structure, SEB forms interactions with both the HLA and TCR, breaking all peptide-TCR contacts in the process. Thus, it is possible for non-peptidic species to form specific interactions with the ternary complex and induce peptide-independent activation of the TCR and the immune response.

It seems unlikely, however, that glycans would improve TCR induction via a specific mechanism (either peptide-dependent or peptide-independent), as glycan identity and location varies among the glycopeptides in this study. Given the known connection between aberrant glycosylation and cancer, it seems more likely that glycans located on residues within and adjacent to the peptide-binding groove serve as a steric barrier that prevents the TCR from associating with the HLA-glycopeptide complex, allowing cells expressing such glycopeptides to evade the immune system and potentially become cancerous.

Identification and Characterization of HLA-DR-Associated Disulfide Bonded Glycopeptides

While sequencing the peptides listed in Table 1, we noticed an interesting phenomenon. ETD generates fragments called c- (N-terminal) and z•- (C-terminal) ions, where the mass difference between members of the same series permits identification of a particular amino acid and assign a peptide sequence. Several ETD spectra had multiple c and z• patterns, suggestive of at least two peptides present in the isolation window. This can occur by one of two mechanisms: (a) co-eluting species of the same mass or (b) disulfide bonded peptides.

After sequencing both peptides present in the ETD spectra, we have determined that the latter is present in both 2048mel and 1363mel, but not in the corresponding EBV cell line. We sequenced a total of 15 peptides containing a disulfide bond (Table 3), all originating from lysosome-associated membrane glycoprotein 1 (LAMP1, Uniprot: P11279), a protein implicated in antigen processing and tumor cell metastasis.³³ The disulfide bond connects C155 and C191, a known disulfide linkage, suggesting that the peptide is presented in this form as opposed to dimerization during sample processing or column elution. Additionally, these peptides contain a variable amount of glycosylation present on Asn165, ranging from a single GlcNAc to GlcNAc₂Fucose₁Mannose₃. To our knowledge, these peptides represent the first known disulfide bonded peptides to be presented by the MHC class II processing pathway.

As the disulfide bonded peptides were exclusively found in melanoma cell lines, we sought to develop a model for their development. Gamma interferon-inducible lysosomal thiol-reductase (GILT) reduces disulfide bonds in the lysosome, facilitating the complete unfolding of proteins before presentation via MHC class II.³⁴ Several studies have shown that human melanoma cells lack this enzyme.³⁵ Further, downregulation of GILT is related to reduced CD4+ T cell recognition;³⁵ however, the mechanism for this is unknown. Thus, we propose a potential mechanism for immune evasion by melanoma cells. With the downregulation of GILT, it is possible that protein disulfide bonds are not completely reduced, and peptides from these proteins are then presented to CD4+ T cells. As the peptides will not fit in TCR geometry, they will not stimulate a response. Future studies will be devoted to understanding GILT expression and its relation to the presentation of disulfide bonded peptides, which may prove promising in understanding immune evasion in cancer.

Materials and Methods

Cultured Cell Lines

HLA genotypes of cultured melanoma cell lines were determined by the NIH Warren Grant Magnuson Clinical Center HLA Laboratory (Bethesda, MD) by using sequence-specific PCR techniques. The cultured plastic-adherent melanoma lines 1102mel, 1363mel, and 2048mel were initiated from enzymatic digests of metastatic lesions. The 2048-EBV cell line was generated from PBMCs according to standard methods and was maintained as a suspension culture. All cell lines were maintained at 37°C, 5% CO₂ in medium consisting of RPMI 1640 plus 10% heat-inactivated FCS, 2 mM L-glutamine, 10 mM HEPES buffer, and antibiotics. Cultures were certified to be free of *Mycoplasma* contamination by PCR.¹⁶

Isolation of HLA-DR-Associated Peptides

To prepare cells for extraction of MHC-peptide complexes, growing cultures were harvested, and spent medium was removed by centrifugation. Cells were washed twice in cold PBS, and dry cell pellets were snap frozen on dry ice and stored at -80 °C for subsequent lysis. Cells (total number: 6.5e8 of 1102mel, 1.4e8 of 2048mel, 3.4e8 of 1363mel, and 9e8 of 2048EBV) were lysed in a solution of 20 mM Tris HCl, pH 8.0; 150 mM NaCl with 1% CHAPS; 1 mM PMSF; 5 g/mL aprotinin; 10 g/mL leupeptin; 10 g/mL pepstatin A; and 1:100 dilutions of phosphatase inhibitor cocktails I and II (Sigma Aldrich),

to prevent potential dephosphorylation of peptides during extraction. The lysate was centrifugated and then run over an Econocolumn (Bio-Rad) containing the pan-HLA-DR-specific mAb L243 bound to recombinant protein. After over-night incubation at 4 °C, peptides were eluted from HLA-DR molecules with 10% acetic acid (HOAc) and separated using a 10-kDa cutoff ULTRAFREE-MC filter (Millipore). Extracts were stored at –80 °C.¹⁶

RP-HPLC-Mass Spectrometry

Samples were taken to dryness and brought up in 0.1% acetic acid. 1×10^7 cell equivalents were directly loaded onto an in-house, packed C18 column (prepared as previously described).³⁶ Briefly, an irregular C18 (5–20 µm diameter) capillary precolumn (360-µm outer diameter, 75-µm inner diameter) was connected to a C18 (5 µm diameter) analytical capillary column (360-µm outer diameter, 50-µm inner diameter) equipped with an electrospray emitter tip. Peptides were eluted by a 90 minute 0–60% B gradient (A: 0.1M acetic acid, B: 70% ACN, 0.1M acetic acid) using an Agilent 1100 HPLC at a flow rate of 60 nL/min. The RP-HPLC elution was electrospray-ionized into an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). Instrument method parameters were as follows: MS1 resolution 60,000 at 200 m/z, scan range 300–1500 m/z, peak width 2 m/z, HCD voltage at 25% collision energy, 50 millisecond ETD reaction time. The instrument selected the most abundant ions with HCD (top speed, 3 seconds) and triggered ETD only when three of six HexNAc fingerprint ions (m/z 204, 186, 168, 144, 138, and 126 $\pm$ 0.2 Da) were detected at >5% relative abundance.

Data Analysis

Peptide sequences were determined by *de novo* manual interpretation of HCD and ETD mass spectra. Data for the glycopeptide sequences reported here were unambiguous. None of the sequences were assigned by software. Approximate copy/cell numbers for each glycopeptide were determined by comparing peak areas of the observed parent ions to the average abundance of angiotensin and vasoactive peptide (100 femtomoles) loaded onto the HPLC column concurrently with sample. Specific functional and intracellular localization information for source proteins was determined from the Uniprot database (www.uniprot.org).

Modeling of Peptide-MHC Complexes and TCR

Modeling was performed using the 2014 Molecular Operating Environment (MOE) software suite.³⁷ Five ternary HLA-peptide-TCR complexes (PDB IDs: 4OZG, 3T0E, 4C56, 4E41, and 1FYT)^{26–30} were superimposed to achieve maximum overlap of the α and β HLA subunits (RMSD C $_{\alpha}$ between 0.45–1.04 Å and 0.54–1.04 Å respectively). When overlaid in this manner, the backbones of the peptide ligands (RMSD C $_{\alpha}$ between 0.48–2.00 Å) have significant overlap, unlike the backbones of the α and β TCR subunits (RMSD C $_{\alpha}$ between 8.08–24.56 Å and 8.46–26.75 Å respectively). The sequences of the three glycopeptides were aligned with those of the crystallized peptides (Table 2). Using this sequence alignment and the crystallized peptide GELIGILNAAKVPAD as a scaffold, the peptide portions of the three glycopeptides were constructed in the binding pocket, and each residue's side chain was rotated to optimize interactions with the adjacent residues of the HLA and TCR

complexes. Following this, the glycans of GlcNAc₂Mannose₃ N478 and GlcNAc₂Fucose₁Mannose₃-N243 of gp120 (PDB ID: 2BF1)³¹ were grafted onto the appropriate residues of the previously constructed peptides to yield the three final model glycopeptides. For each glycopeptide, a conformational search was used to generate conformer libraries in which the modified asparagine residue and appended glycan were allowed to sample conformational space. Individual conformers were then analyzed for their potential to disrupt association of the TCR to the HLA-glycopeptide complex, and the most illustrative conformer for each peptide is shown (Figure 2).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We wish to acknowledge Tracee L. McMiller and Theresa S. Pritchard (Johns Hopkins University School of Medicine) for technical assistance. Additionally, we would like to thank Jane V. Aldrich (University of Florida Department of Medicinal Chemistry) for providing access to Molecular Operating Environment (MOE). This work was supported by a research grant from the Melanoma Research Alliance (to S.L.E, D.F.H., and V.H.E) and by National Institutes of Health Grant AI033993 (to D.F.H.). This work was also supported by USPHS Grants R01 AI20963 and CA134060 (to V.H.E.). ALZ was a recipient of a Sidney Kimmel Scholar award.

References

1. Roche PA, Furuta K. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nat Rev Immunol*. 2015; 4:203–16.
2. Purcell AW I, van Driel R, Gleeson PA. Impact of glycans on T-cell tolerance to glycosylated self-antigens. *Immunol Cell Biol*. 2008; 7:574–9.
3. Engelhard VH, Altrich-Vanlith M, Ostankovitch M, Zarling AL. Post-translational modifications of naturally processed MHC-binding epitopes. *Curr Opin Immunol*. 2006; 18(1):92–7. [PubMed: 16343885]
4. Petersen J, Pursell AW, Rossjohn J. Post-translationally modified T-cell epitopes: immune recognition and immunotherapy. *J Mol Med*. 2009; 87:1045–51. [PubMed: 19763524]
5. Ishioka GY, Lamont AG, Thompson D, Bulbow N, Gaeta FCA, Sette A, Grey HM. MHC interaction and T cell recognition of carbohydrates and glycopeptides. *J Immunol*. 1992; 148(8):2446–51. [PubMed: 1560200]
6. Michaelsson E, Malmstrom V, Reis S, Engstrom A, Burkhardt H, Holmdahl R. T cell recognition of carbohydrates on type II collagen. *J Exp Med*. 1994; 180(2):745–9. [PubMed: 8046350]
7. Michaelsson E, Broddefalk J, Engstrom A, Kihlberg J, Holmdahl R. Antigen processing and presentation of a naturally glycosylated protein elicits major histocompatibility complex class II-restricted, carbohydrate-specific T cells. *Eur J Immunol*. 1996; 26:1906–10. [PubMed: 8765038]
8. Dzhambazov B, Holmdahl M, Yamada H, Lu S, Vestberg M, Holm B, Johnell O, Kihlberg J, Holmdahl R. The major T cell epitope on type II collagen is glycosylated in normal cartilage but modified by arthritis in both rats and humans. *Eur J Immunol*. 2005; 35(2):357–66. [PubMed: 15682451]
9. Mastrangelo A, Colasanti T, Barbati C, Pecani A, Sabatinelli D, Pendolino M, Truglia S, Massaro L, Mancini R, Miranda F, Spinelli FR, Conti F, Alessandri C. The role of posttranslational protein modifications in rheumatological diseases: focus on reumatoid arthritis. *J Immunol Res*. 2015:712490. [PubMed: 26090496]
10. Housseau F, Moorthy A, Langer DA, Robbins PF, Gonzales MI, Topalian SL. N-linked carbohydrates in tyrosinase are required for its recognition by human MHC class II-restricted CD4+ T cells. *Eur J Immunol*. 2001; 31:2690–701. [PubMed: 11536167]

11. Chicz RM, Urban RG, Gorga JC, Vignali DA, Lane WS, Strominger JL. Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *J Exp Med.* 1993; 178(1):27–47. [PubMed: 8315383]
12. Dengjel J, Rammensee HG, Stevanovic S. Glycan side chains on naturally presented MHC class II ligands. *J Mass Spectrom.* 2005; 40(1):100–4. [PubMed: 15643650]
13. Harding CV, Kihlberg J, Elofsson M, Magnusson G, Unanue ER. Glycopeptides bind MHC molecules and elicit specific T cell responses. *J Immunol.* 1993; 151:2419–25. [PubMed: 8360471]
14. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science.* 2009; 324(5930):1029–33. [PubMed: 19460998]
15. Morrish F, Isern N, Sadilek M, Jeffrey M, Hockenbery DM. c-Myc activates multiple metabolic networks to generate substrates for cell-cycle entry. *Oncogene.* 2009; 28(27):2485–2491. [PubMed: 19448666]
16. Depontieu FR, Qian J, Zarling AL, McMiller TL, Salay TM, Norris A, English AM, Shabanowitz J, Engelhard VH, Hunt DF, Topalian SL. Identification of tumor-associated, MHC class I-restricted phosphopeptides as targets for immunotherapy. *PNAS.* 2009; 106(29):12073–8. [PubMed: 19581576]
17. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences.* 2004; 101:9528–9533.
18. Lippolis JD, White FM, Marto JA, Luckey CJ, Shabanowitz J, Hunt DF, Engelhard VH. Analysis of MHC class II antigen processing by quantitation of peptides that constitute nested sets. *J Immunol.* 2002; 169:5089–5097. [PubMed: 12391225]
19. Khong HT, Rosenberg SA. Pre-existing immunity to tyrosinase-related protein (TRP)-2, a new TRP-2 isoform, and the NY-ESO-1 melanoma antigen in a patient with a dramatic response to immunotherapy. *J Immunol.* 2002; 168(2):951–6. [PubMed: 11777994]
20. Schwartz JD, Rowinsky EK, Yousoufian H, Pytowski B, Wu Y. Vascular endothelial growth factor receptor-1 in human cancer. *Cancer.* 2010; 116(4):1027–32. [PubMed: 20127948]
21. Yu JA, Sadaria MR, Meng X, Mitra S, Ao L, Fullerton DA, Weyant MJ. Lung cancer cell invasion and expression of intercellular adhesion molecule-1 (ICAM-1) are attenuated by secretory phospholipase A2 inhibition. *J Thorac Cardiovasc Surg.* 2012; 143(2):405–11. [PubMed: 22248682]
22. Wolfert MA, Jan Boons G. Adaptive immune activation: glycosylation does matter. *Nat Chem Biol.* 2013; 9(12):776–84. [PubMed: 24231619]
23. Queiroz RMD, Carvalho Æ, Dias WB. O-GlcNAcylation: The Sweet Side of the Cancer. *Front Oncol Frontiers in Oncology.* 2014; 4:132.doi: 10.3389/fonc.2014.00132 [PubMed: 24918087]
24. Xu Y, Sette A, Sidney J, Gendler SJ, Franco A. Tumor-associated carbohydrate antigens: a possible avenue for cancer prevention. *Immunology and Cell Biology.* 2005; 83:440–8. [PubMed: 16033540]
25. Engelhard VH, Brickner AG, Zarling AL. Insights into antigen processing gained by direct analysis of the naturally processed class I MHC associated peptide repertoire. *Mol Immunol.* 2002; 39:127–37. [PubMed: 12200045]
26. Petersen K, Monterrat V, Mujico JR, Loh KL, Beringer DX, van Lummel M, Thompson A, Mearin ML, Schweizer J, Kooy-Winkelaar Y, van Bergen J, Drijfhout JW, Kan WT, La Gruta NL, Anderson RP, Reid HH, Koning F, Rossjohn J. T-cell receptor recognition of HLA-DQ2-gliadin complexes associated with celiac disease. *Nat Struct Mol Biol.* 2014; 21(5):480–8. [PubMed: 24777060]
27. Yin Y, Wang XX, Mariuzza RA. Crystal structure of a complete ternary complex of T-cell receptor, peptide-MHC, and CD4. *Proc Natl Acad Sci USA.* 2012; 109(14):5405–10. [PubMed: 22431638]
28. Rodstrom KE, Elbing K, Lindkvist-Petersson K. Structure of the superantigen staphylococcal enterotoxin B in complex with TCR and peptide-MHC demonstrates absence of TCR-peptide contacts. *J Immunol.* 2014; 193(4):1998–2004. [PubMed: 25015819]
29. Deng L, Langley RJ, Brown PH, Xu G, Teng L, Wang Q, Gonzales MI, Callender GC, Nishimura MI, Topalian SL, Mariuzza RA. Structural basis for the recognition of mutant self by a tumor-

- specific, MHC class II-restricted T cell receptor. *Nat Immunol.* 2007; 8(4):398–408. [PubMed: 17334368]
30. Hennecke J, Carfi A, Wiley DC. Structure of a covalently stabilized complex of a human alphabeta T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1. *EMBO J.* 2000; 19(21):5611–24. [PubMed: 11060013]
31. Chen B, Vogan EM, Gong H, Skehel JJ, Wiley DC, Harrison SC. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature.* 2005; 433(7028):834–41. [PubMed: 15729334]
32. Jo S, Lee HS, Skolnick J, Im W. Restricted N-glycan conformational space in the PDB and its implication in glycan structure modeling. *PLoS Comput Biol.* 2013; 9(3):e1002946. [PubMed: 23516343]
33. Agarwal AK, Srinivasan N, Godbole R, More SK, Budnar S, Gude RP, Kalraiya RD. Role of tumor cell surface lysosome-associated membrane protein-1 (LAMP1) and its associated carbohydrates in lung metastasis. *J Cancer Res Clin Oncol.* 2015; 141(9):1563–71. [PubMed: 25614122]
34. Hastings KT. GILT: Shaping the MHC class II restricted peptidome and CD4(+) T cell-mediated immunity. *Front Immunol.* 2013; 4:429. [PubMed: 24409178]
35. Goldstein OG, Hajiaghamohseni LM, Sundaram K, Reddy SV, Hague A. Gamma-IFN-inducible-lysosomal thiol reductase modulates acidic proteases and HLA class II antigen processing in melanoma. *Cancer Immunol Immunother.* 2008; 57(10):1461–70. [PubMed: 18343923]
36. Udeshi ND, Compton PD, Shabanowitz J, Hunt DF, Rose KL. Methods for analyzing peptides and proteins on a chromatographic timescale by electron-transfer dissociation mass spectrometry. *Nat Protoc.* 2008; 3(11):1079–17.
37. Molecular Operating Environment (MOE), 2014.09. Chemical Computing Group Inc; 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada: 2015. H3A 2R7

Figure 1. Glycoforms detected on MHC class II eluted peptides

Seventeen glycoforms were detected on peptides. Numbers listed in this figure are listed under “glycosylation” in table 1. Note that the dissociation method used here provides amino acid sequence information but does not fragment the glycans themselves, so the data reported here does not define the glycan structures. Branching patterns are based on known glycan structures along with sequential neutral losses from mass spectra.

Figure 2. Overlay of modeled HLA-glycopeptide-TCR ternary complexes

The glycans of glycopeptides 1 (orange), 2 (yellow), and 3 (green) can adopt conformations in which they interact with the TCR (pink/purple) but not the HLA (cyan/blue). The glycans flank the residues comprising the peptide binding groove and could affect antigen recognition by the TCRs, which vary among themselves in structure.

Table 1

List of MHC class II associated glycopeptides

Eighty glycopeptides are listed above. Protein sources were determined by searching the peptide sequence against the nr databases for human proteins (www.ncbi.nlm.nih.gov/BLAST). Relative abundance key: nd, not detected, + 1–5 copies/cell, ++ 6–50 copies/cell, +++ 50–150 copies/cell, and ++++ >150 copies per cell. Additional information for the peptides can be found in SI Table 2.

	Source Protein	Peptide	Glyco form	1102 Mel	1363 Mel	2048 Mel	2048 EBV
1	Alpha-2-macroglobulin	859AVTPKSLGNVn869	4	nd	nd	++	nd
2	Cation-independent mannose-6-phosphate receptor	123IDPRHGNLYDLKPLGLnDT1248	8	nd	++	nd	nd
3	CD109 antigen	67SnLTVSVLEAGVFKEK82	8	nd	+++	nd	nd
4	CD97 antigen	94SPGYEPVSGAKTFKn108	6	nd	++	nd	nd
5			8	nd	++	nd	nd
6		94SPGYEPVSGAKTFKnE109	6	nd	+	nd	nd
7		94SPGYEPVSGAKTFKnE110	8	nd	++	nd	nd
8	Cholinesterase	32IIIATKNGKVRGMnLTVFGGTVT54	13	nd	nd	++	nd
9			14	nd	nd	+++	nd
10			15	nd	nd	++	nd
11		32IIATKNGKVRGMnLTVFGGTVTAF56	13	nd	nd	++	nd
12			14	nd	nd	+++	nd
13			15	nd	nd	+++	nd
14	Collagen alpha-1(XVIII) chain	1555HSSYVHLRPARPISPP1570	2	nd	++	nd	nd
15	EF-hand calcium-binding domain-containing protein 14	334SATLKRQSLDQVTnRT349	6	nd	nd	++	nd
16	Fibronectin	2006HGPEILDVPSTVQKp2022	1	nd	nd	+++	nd
17	Glia-derived nexin	106DIVTVANAVFVKnAS120	8	nd	++	nd	nd
18		105KDIVTVANAVFVKnAS120	8	++	++	nd	nd
19			10	nd	++	nd	nd
20		105KDIVTVANAVFVKnASE121	8	nd	++	nd	nd
21		104NKDIVTVANAVFVKnAS120	8	nd	+++	nd	nd
22		103KNKDIVTVANAVFVKnAS120	8	nd	+	nd	nd
23		103KNKDIVTVANAVFVKnASE121	6	nd	nd	++	nd
24			8	nd	nd	++	nd

	Source Protein	Peptide	Glyco form	1102 Mel	1363 Mel	2048 Mel	2048 EBV
25		102KKNKD ^{IV} TVANAVFV ^K nAS120	8	nd	+	nd	nd
26		102KKNKD ^{IV} TVANAVFV ^K nASE121	6	nd	nd	++	nd
27			8	nd	nd	++	nd
28	HLA class II DP alpha 1 chain	90GLANIA ^{IL} NNNLNTLIQRSnHTQ112	6	nd	nd	nd	++
29		91LANIA ^{IL} NNNLNTLIQRSnHTQ112	8	nd	nd	nd	++
30		91LANIA ^{IL} NNNLNTLIQRSnHTQ112	6	nd	nd	nd	++
31		91LANIA ^{IL} NNNLNTLIQRSnHT111	6	nd	nd	nd	++
32		94IA ^{IL} NNNLNTLIQRSnHTQ112	6	nd	nd	nd	++
33		94IA ^{IL} NNNLNTLIQRSnHT111	6	nd	nd	nd	++
34	Intercellular adhesion molecule 1	385nQTREL ^{RVLYG} PRLDER401	17	nd	++	++	nd
35		384KnQTREL ^{RVLYG} PRLDER401	17	nd	++	++	nd
36		384KnQTREL ^{RVLYG} PRLDERD402	17	nd	+	++	nd
37		382HHKnQTREL ^{RVLYG} PRLDERD402	17	nd	++	+++	nd
38	Interleukin 6 receptor subunit beta	208TSDHINFD ^{PVYKV} KPNPPHn227	8	++	nd	nd	nd
39	Laminin subunit gamma-1	1172DPNnMTLLAE ^{EARKL} AER1189	8	nd	++	nd	nd
40		1172DPNnMTLLAE ^{EARKL} AERH1190	6	nd	nd	++	nd
41			8	nd	++	nd	nd
42	L-dopachrome tautomerase	160TQHWLGLLGPnGTQPQ175	8	nd	+++	nd	nd
43		228ERDL ^{QRLIGn} ESFA241	13	nd	++	nd	nd
44			14	nd	+	nd	nd
45		228ERDL ^{QRLIGn} ESFALPY244	12	nd	++	nd	nd
46			13	nd	++	nd	nd
47			14	nd	++	nd	nd
48	Lysosome membrane protein 2	213YVFLTGEDSYLn224	4	nd	++	nd	nd
49			9	nd	++	nd	nd
50	Lysosome-associated membrane glycoprotein 1	246RKDnT ^{IV} TRLLNINPnKT263	6	nd	nd	++	nd
51		246RKDnT ^{IV} TRLLNINPnKT263	8				
			6	nd	nd	++	nd
			11				

	Source Protein	Peptide	Glyco form	1102 Mel	1363 Mel	2048 Mel	2048 EBV
52		246RKDnTTIVTRLNINPhKT263	6	nd	nd	++	nd
			12				
53		321AnGSLRALQATVGNsYK337	1	nd	++	++	nd
54	N-acetylglucosamine-6-sulfatase	180KYYnYTLsINGKARKHG196	8	+++	nd	nd	nd
55			6	++	nd	nd	nd
56		180KYYnYTLsINGKARKH195	8	+++	nd	nd	nd
57			6	++	nd	nd	nd
58		181YYnYTLsINGKARKH195	8	++	nd	nd	nd
59		180KYYnYTLsINGKARK194	8	++	nd	nd	nd
60			11	+	nd	nd	nd
61		180KYYnYTLsINGKAR193	8	++	nd	nd	nd
62	Neural cell adhesion molecule L1-like protein	79TDHRIPSaNSGTFER93	8	nd	nd	++	nd
63	Neurosecretory protein VGF	85VLLQALDRPAsPP97	2	nd	nd	++++	nd
64		85VLLQALDRPAsPPA98	3	nd	nd	++	nd
65	Platelet glycoprotein 4	221ISKVAIDITYKGRn235	6	nd	nd	+	nd
66	Prosaposin	214TnSTFVQALVEHVKEE229	4	nd	++	nd	nd
67			6	nd	nd	++++	++
68			8	nd	nd	++++	nd
69		213RTnSTFVQALVEHVKEE228	6	nd	++	++	nd
70			8	nd	++	nd	nd
71		213RTnSTFVQALVEHVKEE229	6	nd	++	++	nd
72			8	nd	+++	++	nd
73		212VRTnSTFVQALVEHVKEE228	8	nd	++	nd	nd
74		212VRTnSTFVQALVEHVKEE229	6	nd	+++	nd	nd
75			8	nd	+++	nd	nd
76	Scavenger receptor class B member 1	299YRFVAPKTLFAnGSiYPP316	16	nd	++	nd	nd
77	Sialomucin core protein 24 (MUC24)	125KPtVQFSPSTtSK137	2	nd	nd	nd	+
			2	nd			
78	Vascular endothelial growth factor receptor 1	236TPRPVKLLRGHTLVLn251	7	nd	++	++	nd

	Source Protein	Peptide	Glyco form	1102 Mel	1363 Mel	2048 Mel	2048 EBV
79			9	nd	++	++	nd
80			11	nd	++	++	nd

Table 2

Sequence alignment of peptides used in modeling studies

Peptides in HLA-peptide-TCR crystal structures are either complexed with HLA-DQ (4OZG and 3T0E) or HLA-DR (4C56, 4E41, and 1FYT), with small deviations between Ca atoms of the peptide backbones (RMSD). Glycopeptide sequences were aligned to crystallographic peptide sequences using Immune Epitope Database and Analysis Resource. Modified asparagine residues lower case Predicted nine-residue binding motif is bolded.

Identity (PDB ID)	Sequence										RMSD (in Å)											
	I		5		10		15		40ZG	3TOE	4C56	4E41	1FYT									
40ZG	.	.	A	P	Q	P	E	L	P	Y	P	Q	P	G	S	.	.	-	1.78	1.47	1.24	1.30
3TOE	.	F	S	W	G	A	E	G	Q	R	P	G	F	G	S	G	G	-	1.50	2.00	1.46	
4C56	.	P	K	Y	V	K	Q	N	T	L	K	L	A	T	.	.	1.47	1.50	-	1.01	0.48	
4E41	G	E	L	I	G	I	L	N	A	A	K	V	P	A	D	.	1.24	1.62	1.01	-	0.95	
1FYT	.	P	K	Y	V	K	Q	N	T	L	K	L	A	T	.	.	1.30	1.46	0.48	0.95	-	
Glycopeptide 1	T	P	R	P	V	K	L	L	R	G	H	T	L	V	L	n	.	-	-	-	-	
Glycopeptide 2	S	n	L	T	V	S	V	L	E	A	E	G	V	F	E	K	.	-	-	-	-	
Glycopeptide 3	D	n	T	T	V	T	R	L	L	N	I	N	P	n	K	T	.	-	-	-	-	

Table 3
List of MHC class II associated disulfide-bonded glycopeptides

Listed are fifteen disulfide-bonded glycopeptides eluted from two melanoma cell lines (1363mel and 2048mel). Peptides originate from lysosome-associated membrane protein 1 (LAMP1) and contain known disulfide linkage between C155 and C191. IEDB predicts the nine-residue binding motif to be YRCVSGTQV, suggesting that the disulfide linkage is within the TCR binding pocket. Relative abundance key: nd, not detected, + 1–5 copies/cell, ++ 6–50 copies/cell, +++ 50–150 copies/cell, and ++++ >150 copies per cell. N-linked glycosylation is indicated with a lower case n, oxidized methionine with a lower case m.

Peptide	Glycoform	1363 Mel	2048 Mel
DIDKKYRCVSGTQVHMNn ETRCEQ	8	+	nd
DIDKKYRCVSGTQVHMNn GETRCEQ	7	++	nd
DIDKKYRCVSGTQVHmNn GETRCEQ	6	+	nd
DIDKKYRCVSGTQVHmNnV GETRCEQ	6	+	nd
IDKKYRCVSGTQVHMNn C	5	++	nd
	7	+	nd
	8	+	nd
IDKKYRCVSGTQVHMNn CE	4	nd	+
	5	nd	+
IDKKYRCVSGTQVHMNn ETRCEQ	4	++	nd
	5	++	nd
	6	+	nd
	8	++	nd
IDKKYRCVSGTQVHmNn ETRCEQ	4	+	nd
	5	+	nd
	8	+	nd
IDKKYRCVSGTQVHMNn ETRCEQD	4	+	nd
	5	++	+
	6	+	nd
IDKKYRCVSGTQVHmNn ETRCEQD	5	+	nd
IDKKYRCVSGTQVHMNn GETRCEQ	4	+	nd
	5	++	+
	7	+	nd

Peptide	Glycoform	1363 Mel	2048 Mel
	8	++	nd
IDKKYRCVSGTQVHmNn	4	+	nd
GETRCEQ	5	+	nd
	8	+	nd
IDKKYRCVSGTQVHmNnV	4	+	nd
GETRCEQ			
IDKKYRCVSGTQVHMNnVT	8	+	nd
GETRCEQ			
IDKKYRCVSGTQVHMNn	6	+	nd
GETRCEQD			