2018-05-15

An Examination of MHC, Peptide, and TCR Interactions

Peter Trenh
University of Massachusetts Medical School

AN EXAMINATION OF MHC, PEPTIDE, AND TCR INTERACTIONS

A Dissertation Presented

By

PETER TRENH

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

MAY 15TH, 2018

IMMUNOLOGY AND MICROBIOLOGY PROGRAM

AN EXAMINATION OF MHC, PEPTIDE, AND TCR INTERACTIONS

A Dissertation Presented by
PETER TRENH

The signatures of the Dissertation Defense Committee signify completion and approval as to style and content of the Dissertation

_____
Lawrence J. Stern, Ph.D., Thesis Advisor

_____
Richard Dutton, Ph.D., Member of Committee

_____
Sharone Green, M.D., Member of Committee

_____
Celia A. Schiffer, Ph.D., Member of Committee

_____
Scott C. Garman, Ph.D., External Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation committee

_____
Stuart Levitz, M.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school

_____
Mary Ellen Lane, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

Immunology and Microbiology Program
May 15th, 2018

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Lawrence Stern, for guiding me through many difficult projects and for fostering an environment that encourages independent learning. I will feel fortunate if I can retain one percent of the breadth and detail of science that you can retain. Thank you for showing me how to approach science and how to balance that with life.

I would also like to thank my TRAC and DEC committees for their time and insight. Every comment and advice spurred new avenues in my research and I am grateful for your expertise.

I would also like to thank the Stern lab's past and present members. Our conversations ran the gamut from abstract to absurd. Thank you for the memories and the calories.

Lastly, I would like to thank my wife, Alison Bright, for being my anchor and voice of reason. Thank you for always reminding me about the light at the end of the tunnel and for dealing with my idealisms. Watching paint dry would have been more exciting than waiting for me to finish my dissertation. Thank you for the endless patience, love, and support.

ABSTRACT

T cell receptors (TCR) bind to peptides from various sources on MHC (Major Histocompatibility Complex) molecules. A long-standing goal in the field is to understand the mechanisms of MHC-peptide exchange and MHC-TCR interactions. Here, I present work from three uniquely different systems that address the following: HLA-DR1 conformational stability, self-tolerant mechanisms of TCRs isolated from self-reactive TCR transgenic mice, and TCR cross-reactivity mechanisms between LCMV and VV.

First, I present a crystal structure of HLA-DR1 in complex with A1L9 peptide, a peptide with two amino acid substitutions from the parental peptide. The singly substituted A1 peptide, which has a pocket 1 alanine substitution, decreases intrinsic half-life between MHC-peptide and increases susceptibility to HLA-DM mediated peptide exchange. This data agrees with previous models of HLA-DM-mediated peptide exchange in which the major determinant is located at the HLA-DR1 pocket 1. However, the L9 substituted peptide, which has a pocket 9 leucine substitution, displays the opposite phenotype: increased intrinsic half-life and decreased HLA-DM susceptibility. The crystal structure presented here shows that HLA-DR1 in complex with a doubly substituted peptide, A1L9, is in the same conformation as HLA-DR1 with the wild-type peptide, demonstrating that pocket 9 residues can rescue pocket 1 residue binding deficiencies and that HLA-DR1 stability is determined by amino acids along the peptide, not only at pocket 1.

Next, I present crystal structures of two self-tolerant TCRs in complex with IAb-3K pMHC. To elucidate molecular mechanism for self-reactivity and self-tolerance, the TCRs J809.B5 and 14.C6 are compared to each other and its parental self-reactive TCR, YAe-62.8. In comparison to YAe-62.8, J809.B5 interacts with the same pMHC, but utilizes more peptide specific interactions, a mechanism that may distinguish self-reactive receptors from self-tolerant receptors. Additionally, the crystal structure of 14.C6 TCR, which bears a different CDR3α sequence from J809.B5, demonstrates that CDR3 sequences can modulate interactions of germline encoded CDR1 and CDR2 loops. Together, these results highlight that in addition to CDR3 VDJ recombination, diversity is generated in the mature TCR repertoire by differential chain pairing, either of which can affect the interactions of germline encoded CDR loops.

Next, I present a detailed analysis of cross-reactive TCRs between K^b-GP34 and K^b-A11R. The mature LCMV-immune repertoire was analyzed by DNA deep sequencing of TCRβ CDR3 sequences, which led to the identification of new cross-reactive sequence motifs. Cross-reactive sequence motifs varied by each Vβ gene, suggesting a role of CDR1, CDR2, and CDR3 loop interplay in cross-reactivity.

Lastly, I present the crystal structures of a GP34/A11R cross-reactive TCR in complex with both K^b-GP34 and K^b-A11R. Analysis of the crystal structures revealed that the two complexes are largely the same, despite differences in peptide sequences. Surprisingly,

the TCR to peptide interactions were dominated by three out of eight peptide side-chains. Cross-reactivity between these two complexes is likely due to a large amount of interactions from TCR to MHC compared to interactions of TCR to peptide. We note two unique MHC-peptide interactions that may allow K^b to be an allele prone to cross-reactivity. The first is an interaction at the C-terminus of the A11R peptide which pulls A11R P7 asparagine away from TCR interactions. The second interaction is from an arginine at position 155, which sits at the interface between TCRα and TCRβ , and contributes the most buried surface area in the interaction interface. Because K^b's arginine 155 is a long side chain that hydrogen bonds with the peptide backbone, and is also at the center of the TCR-peptide interface, GP34 and A11R peptide sequence differences may be occluded from TCR discrimination by K^b presentation.

The data presented in this dissertation demonstrate that interactions between MHC-peptide and MHC-TCR act harmoniously and coopertively, whereby proximal interactions are affected by interactions elsewhere. While previous models of HLA-DR/HLA-DM interactions demonstrate the importance of interactions at HLA-DR1 pocket 1, I showed that pocket 9 also contributes to HLA-DR stability and therefore, HLA-DM susceptibility. I also showed that TCR CDR3 loop sequences affect germline CDR1/CDR2 loop interactions and vice versa. Lastly, I showed that allele specific MHC side chain interactions with the bound peptide influence TCR ligand binding and hence, TCR cross-reactivity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

The work presented here represents data that is published as part of collaborations with other investigators (Chapters II and III), and work in preparation for submission (Chapters IV and V). Except for Chapters I and VI, a detailed introduction will precede the results to introduce the specific research problem addressed. Chapters II and III contain a synopsis of co-authored manuscripts to highlight my contributions with respect to each publication.

CHAPTER I

GENERAL INTRODUCTION

The adaptive branch of the immune system protects animals against pathogens, which evolve at a faster rate than humans. The unpredictable variability of pathogens is met with a high degree of variability from the adaptive immune system. While antibodies are directed against soluble antigens, T cell receptors (TCRs) are directed towards cellular antigens. Antigenic peptides are presented on the cell surface while bound by Major Histocompatiblity Complex (MHC) proteins and the MHC-peptide complex is recognized by TCRs. The field of adaptive immunity has benefitted immensely from crystal structures of MHC and TCR complexes; however, few examples represent TCR cross-reactivity, a general phenomenon of TCRs. Thus it is often difficult to formulate mechanistic models that encompass protein dynamics from a few protein structures. This thesis will examine structural features of novel MHC and TCR interactions, providing additional data points for a more nuanced model, and ultimately demonstrating conformational co-cooperativity between different antigen presentation and recognition components.

Antigen Presentation by MHC-I and MHC-II

The MHC is a locus on the genome involved in graft transplantation compatibility. MHC proteins were originally thought to be related to, but different than, blood antigens A, B, and O, and were initially discovered as “antigen II” (Klein, 1986). In mice, MHC genes

are located on chromosome 17 and named “H-2”, from the “histocompatibility” of antigen II. In humans, the MHC genes are located on chromosome 6 and referred to as HLA (Human Leukocyte Antigen), from the discovery of leukocyte agglutination between mismatched donors. While many immunologically relevant genes are located in the MHC locus, some are structurally unrelated to the MHC-fold. On the contrary, other proteins also adopt the MHC-fold, such as the neonatal Fc Receptor and ZAG-1 (Zn-alpha2-glycoprotein), but are not related to peptide presentation. (Burmeister et al., 1994; Sánchez, 1999; Sánchez et al., 1997) (Figure I-1). For the purposes of this dissertation, I will refer to MHC proteins as MHC genes with the same MHC-fold and are involved in peptide presentation,

MHC proteins are an evolutionarily conserved component of the adaptive immune system. Along with other proteins from the adaptive immune system, they are found in jawed chordates and are believed to have evolved after the phylogenetic split from jawless lampreys and hagfish (Danchin et al., 2004). Despite evolutionary conservation, genetic diversity in MHC proteins is important in buffering against pathogenic species and preventing population epidemics (Sommer, 2005). In comparison to the genetically diverse outbred populations, inbred populations are more susceptible to epidemics by pathogens. Pathogens typically have a fast rate of evolution and can evolve new immune evasion tactics. MHC genes are the most polymorphic in the human genome (Hu et al., 2002) which allows them to keep up with the fast pace of pathogen evolution and makes them a key mechanism to combat pathogens.

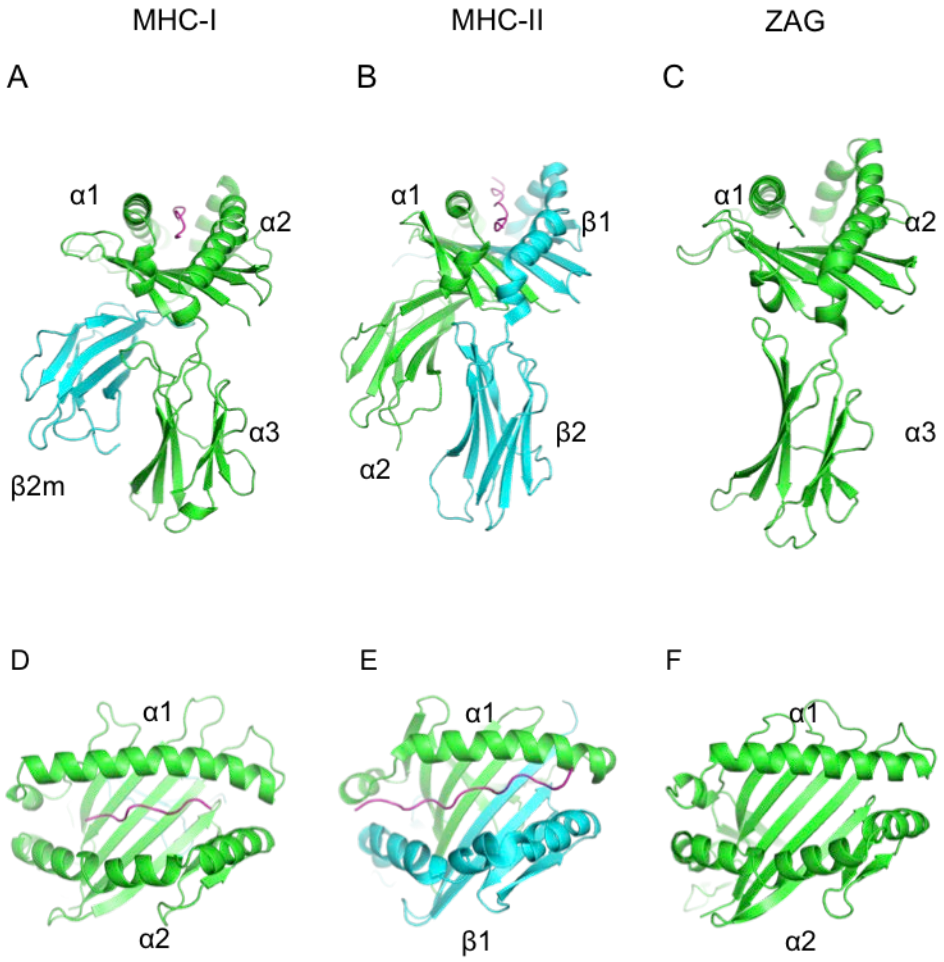

Figure I-1 Comparison of Proteins with MHC-folds.
(A) HLA-A2 is a MHC-I protein that is encoded by two polypeptide chains: (α and β2m or heavy chain and light chain respectively). (PDB: 1AO7). (B) HLA-DR1 is a MHC-II that is encoded by two polypeptide chains: α and β. (PDB: 1DLH). (C) ZAG is a protein with an MHC-fold that does not bind to peptides. (PDB: 1ZAG). (D, E, F) "Top down" view of A, B, and C respectively, with N-terminus (pink) of the peptide to the left. Domains are labeled accordingly with respective nomenclature.

Broadly, MHC proteins are divided into two classes with similar protein folds: MHC-I and MHC-II. The MHC protein fold consists of three distinct domains: a peptide-binding domain, which consists of two α-helices atop eight antiparallel β-strands, along with two Ig (immunoglobulin) domains. As type-I transmembrane proteins, the C-terminal Ig domains are proximal to the cell membrane, while the peptide-binding domain (α-helices and β-strands) is most distal to the cell membrane. Polymorphisms between different MHC alleles are primarily located at peptide binding residues, thus, different MHC alleles have differential peptide binding preferences (Bjorkman and Parham, 1990). Two of the first MHC-I crystal structures solved were H-2K^{b} and HLA-A2. The first structures revealed that peptides bind between the two α-helices, with a number of side chains buried. Comparison between the two complexes led to the identification of pockets that are labeled from A to F — a nomenclature borrowed from identified binding pockets in enzymes. The identification of MHC pockets led to the idea that differential usage of deep and shallow pockets allows for the differential binding of a broad range of peptides. Different MHC alleles may use different combinations of side chains to make up the composition of peptide binding pockets (Matsumura et al., 1992).

The pockets that anchor peptides on the MHC are typically hydrophobic residues, aromatic residues, and sometimes, charged residues. To illustrate the effect of pocket residues, consider the MHC alleles: HLA-B44 and HLA-B27. In HLA-B44 alleles, the MHC A-pocket residue that interacts with the second amino acid of the peptide (P2) is primarily a positively charged lysine. The amino acid typically found at the P2 of

peptides eluted from HLA-B44 alleles is a negatively charged glutamic acid. On the contrary, the corresponding residue in HLA-B27 is negatively charged glutamic acid, with a preference of peptides that contain positively charged arginine, lysine, and histidine at P2 (Rammensee, 1995; Rammensee et al., 1993; Zhang et al., 1998).

Comparing the two mouse MHC-I alleles, H-2K^b (K^b) and H-2D^b (D^b), illustrates that the amino acid composition of MHC-I specificity pockets affects more than just the selection of peptide side chains (Young et al., 1994). While K^b has an octamer peptide length preference, D^b alleles preferentially bind to nonamer peptides. The MHC-I amino acid separating the C-pocket and F-pocket is residue number 73, which is a phenylalanine in K^b and a tryptophan in D^b. The tryptophan in D^b forms a ridge in the peptide binding-groove, which moves the peptide backbone away from the MHC and makes room for a longer peptide to bind in the peptide-binding groove (Figure I-2). Examples of different peptide backbone torsions have also been observed in H-2L^d, H-2K^d, and H-2D^d alleles (Achour et al., 1998; Mitaksov and Fremont, 2006). Thus, differences in MHC alleles can result in changes of peptide presentation which allows different MHC alleles to bind peptides of varying amino acids, length, and shape.

In a strikingly parallel fashion, crystal structures of MHC-II proteins were also determined and their peptide specificity pockets also identified (Brown et al., 1993; Stern and Wiley, 1994). In contrast to MHC-I, MHC-II proteins present longer peptides which

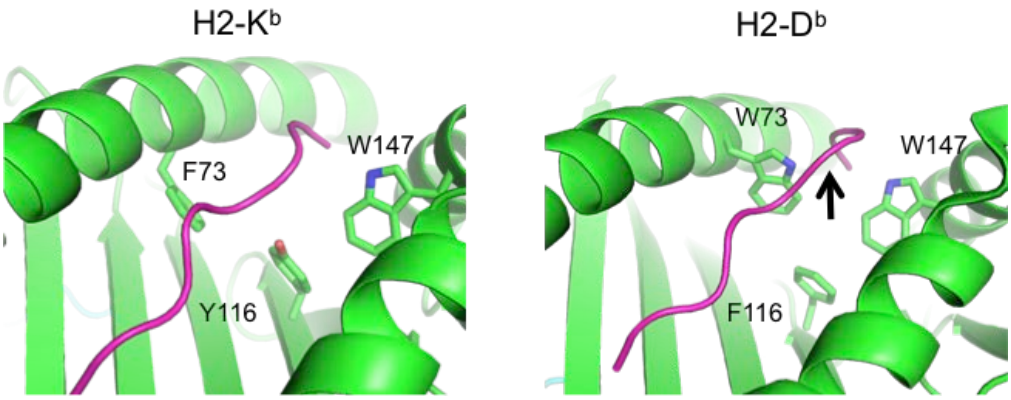

Figure I-2 MHC Allelic Side Chains Influences Peptide Backbone Torsions
(A, B) VSV peptide in H2-K^b is flatter than H2-D^b-NP(366-374) because D^b Side chains form a ridge and torsions the peptide backbone. (PDBID: 2VAA, 1HOC)

extend beyond the anchor pockets. Besides peptide length, a functional difference between MHC-I and MHC-II proteins are the cell types that express each respective protein, and the cell types which recognize the presented peptides. MHC-I proteins are thought to be expressed on virtually all nucleated cells and present peptides to CD8+ T cells, while MHC-II proteins are expressed on professional antigen presenting cells (APCs) and present peptides to CD4+ T cells. The role of CD4 and CD8 co-receptors in driving T cell specificity for MHC-I versus MHC-II is unclear (Van Laethem et al., 2012), but it is evident from crystal structures and binding studies which use soluble recombinant protein that MHC-TCR binding does not depend on co-receptor binding.

Interestingly, MHC-I and MHC-II proteins acquire peptides from different pathways and cellular compartments. The simplified role of MHC-I proteins is to sample and report cytosolic peptides to the immune system (Vyas et al., 2008). Short peptide products that have been degraded by the proteosome are transported to the endoplasmic reticulum where they are loaded onto MHC-I molecules with the aid of accessory proteins such as calreticulin and tapasin. Peptide-MHC-I complexes are then trafficked through the Golgi network to the cell surface. In contrast, MHC-II proteins sample peptides from endocytic compartments. Peptides are loaded onto MHC-II proteins in a late endosomal compartment, termed MHC class II compartment (MIIC) (Neefjes et al., 2011; Stern et al., 2006). Because endocytic compartments typically contain exogenous proteins from endocytosis, MHC-II proteins typically present extracellular antigens. However, evidence of cross-presentation suggest that dendritic cells can also present extracellular

antigens via the MHC-I pathway (Joffre et al., 2012). Ultimately, antigenic peptides presented by MHC proteins are detected by T cell receptors (TCRs). A thorough understanding of how peptides are selected and loaded onto MHC proteins is important in deciphering which antigens are recognized by TCRs.

Newly synthesized MHC-II proteins are folded with a chaperone protein called CLIP. In humans, endosomal peptides are exchanged on MHC-II proteins by a non-classical (non-polymorphic and non-peptide binding) MHC molecule, HLA-DM (DM) (Painter and Stern, 2012). In the case of DM-mediated peptide exchange with HLA-DR1 (DR1), CLIP is exchanged for an optimal peptide with an increased half-life. Thus, DM catalyzes exchange of weakly associating peptides for more stable complexes. Furthermore, peptides with longer DM-mediated half-lives were found to be more immunogenic than peptides with shorter DM-mediated half-lives (Yin et al., 2012).

DM plays an important role in shaping the immunogenic peptide repertoire and understanding the molecular mechanisms by which DM functions on DR1 will help us understand how immunogenic peptides are selected. Chapter II of this dissertation explores the molecular mechanisms of DM-mediated peptide exchange by examining peptide sequences that contribute to DR1 conformational lability.

Antigen Recognition by T cell receptors

The role of MHC proteins in the immune system is peptide presentation to T cells. T cells that recognize antigenic peptide-MHC complexes can initiate immune effector responses such as recruitment of other immune cells or direct killing of the target cell. Thus, MHC polymorphisms allow the immune system at the population level to recognize a wider array of pathogenic antigens. Surprisingly, genes within the MHC locus are often correlated with increased susceptibility to infectious diseases and autoimmune disorders (Trowsdale and Knight, 2013). In rheumatoid arthritis and HIV-1 infection, the etiological genetic associations of MHC to disease are located at peptide binding residues (Matzaraki et al., 2017; Pereyra et al., 2010). The correlations between MHC allele and disease can be caused by an allele's increased preference to bind self-peptides or alternatively, less preference for binding antigenic viral peptides.

A MHC allele's peptide preference ultimately affects selection of mature T cells, the downstream effectors of antigen presentation. T cells are effective at distinguishing self-peptides from foreign peptides because immature T cells undergo two distinct phases of selection. Since self-proteins are always degraded and turned over, self-peptides are always presented on the cell surface by MHC. Immature T cells in the thymus are selected for tonic interactions with MHC first by a positive selection phase where rearranged T cells that do not interact with self-pMHCs die by neglect. T cells are then selected by a negative selection phase during which T cells with strong self-peptide interactions are eliminated via apoptosis (Klein et al., 2014). This double selection education mechanism ensures that the mature T cells which exit the thymus and move to

the periphery are less reactive to MHCs presenting self-peptides and presumably highly reactive to foreign peptides. Taken together, an individual's set of MHC alleles, or haplotype, plays an important role in determining which antigens T cells are reactive against because T cells are educated by their own self-peptide repertoire.

Extrinsic changes to one's self-peptide repertoire can lead to unexpected alloreactive T cell responses. For example, HIV-1 positive individuals who are undergoing anti-retroviral treatment with abacavir, and also express the HLA-B*57:01 allele, develop a condition called abacavir hypersensitivity syndrome; a condition similar to an alloreactive T cell response. Individuals with similar alleles, such as HLA-B*57:02, B*57:03, B*57:11, or B*58:01, do not develop the same symptoms. Careful biophysical characterization and several structural studies determined that abacavir binds within the F pocket of the peptide binding groove in HLA-B*57:01 individuals, therefore changing the specificity of the self-peptide repertoire (Illing et al., 2012; Ostrov et al., 2012). Circulating T cells from these individuals were selected on a peptide repertoire without abacavir. During abacavir treatment, abacavir binds to HLA-B*57:01, altering the preferences of peptides presented, and previously matured circulating T cells recognize these peptide-MHC complexes as foreign peptides. Thus, the self-peptide repertoire, which is dependent on abacavir treatment or an individual's MHC haplotype, directly contributes to the selection of an individual's T cell receptor repertoire.

T cells are activated by binding to peptide-MHCs (pMHC), but the biophysical determinants in TCR ligation remain to be elucidated. From >30 unique pMHC-TCR complexes with MHC-I and >20 TCR-MHC-II complexes, it is evident that TCRs utilize a wide variety of mechanisms to bind pMHC proteins. The seemingly conserved binding mechanism is a diagonal binding mode where TCRα chain contacts MHC-I α2 or MHC-IIβ, and TCRβ chain contacts MHC-I α1 or MHC-IIα (Bridgeman et al., 2012; Rudolph et al., 2006a). However, as has been the case with most "rules" regarding MHC-TCR interactions, exceptions arise. A recent study determined the crystal structure of a TCR binding to pMHC in a reversed polarity compared to conventional pMHC-TCR complexes (Beringer et al., 2015).

The TCR protein consists of four Ig domains: two Ig domains from TCR alpha and two Ig domains from TCR beta. Each TCR chain also has a variable domain and a constant domain. TCR binding to MHCs is mediated through TCR complementarity determining region (CDR) loops located in the variable domain. Each TCR chain (α + β) is comprised of three CDR loops each (CDR1, CDR2, and CDR3). While CDR1 and CDR2 loops are germ-line encoded, CDR3 loops are randomly generated by a process called V(D)J recombination. The immense diversity of TCRs is largely due to V(D)J recombination where one V gene segment (out of >70 Vα or >50 Vβ) combines with one D gene segment (2 Dβ and absent in α), and one J segment (out of >60 Jα or >10 Jβ) (Calis and Rosenberg, 2014). Nucleotides are also added or deleted between each joined

segment, creating another layer of diversity. Lastly, TCRα and TCRβ chain pairing may also contribute to the diversity or restriction of TCRs.

TCR contacts on MHC proteins are more conserved than peptide contacts (Bjorkman and Parham, 1990). This observation suggests that MHC proteins may have evolved to combat evolving pathogens by presenting different peptides on different alleles, while preserving functional MHC-TCR interactions. In support of this co-evolution hypothesis, TCRs use the same set of genes to bind MHC-I, MHC-II, and non-classical MHC-I molecules (Kitamura et al., 1997; Treiner et al., 2003). Another compelling hypothesis regarding MHC restriction is that germ-line encoded CDR1 and CDR2 loops bind to MHCα helices by conserved pairwise interactions, while diverse CDR3 loops are sensitive to peptide sequence differences (Garcia et al., 2009a; Marrack et al., 2008). However, comparison of many MHC-TCR crystal complexes do not reveal conserved patterns of interactions between germ-line encoded regions of TCR and MHC (Baker et al., 2012).

Of interest is how a limited set of TCR genes and their components can affect binding to MHC and peptide recognition. Chapter III of this dissertation will explore how TCR α-TCR β V gene pairing and CDR3 loop sequences affect pMHC ligand binding.

Cross-reactivity and Heterologous Immunity

Based on the lengths of peptides presented by MHC and 20 proteogenic amino acids, it is estimated that $>10^{15}$ possible peptides can be presented (Davis and Bjorkman, 1988; Sewell, 2012), albeit not every possible peptide combination will bind an MHC molecule. While an effective TCR repertoire should ideally respond to any potential foreign peptide antigen, the number of single positive CD4+/CD8+ T cells in a mouse are on the order of 10^7 cells (Sawicka et al., 2014), a number much smaller than the number of possible peptides. If a mouse were to have an immune system agreeing with clonal selection theory, where one TCR recognizes one cognate peptide, it would have 70-fold more T cells (Mason, 1998). Thus, an alternate model to clonal selection theory is that TCRs are inherently cross-reactive in order to recognize the milieu of potential foreign antigens.

Several models of ligand binding have been proposed based on structural studies of MHC-TCR complexes (Yin and Mariuzza, 2009). While the structures of many MHC-TCR complexes have been determined, few capture the pervasive cross-reactive nature of TCRs: the same TCR ligated to different peptide-MHC complexes. Mechanistic models of these cross-reactive MHC-TCR complexes demonstrate the plasticity and immense structural adaptations between MHC, peptide, and TCRs. In the most simplistic mechanism of cross-reactivity, various pMHC complexes present a similar surface for TCR docking. For example, LC13 TCR docks to three pMHCs identically: HLA-B*0801 with an EBV peptide and HLA-B*4405 with an ABCD3 peptide or mimotope (Kjer-Nielsen et al., 2003; Macdonald et al., 2009). The three complexes overlay identically, which suggests that pMHC complexes with different peptides can mimic each other's

surfaces. A second mechanism of TCR cross-reactivity is the case of 2C TCR, which was solved in complex with K^b and the dEV8 peptide, as well as L^d with the QL9 peptide (Colf et al., 2007). Between the two complexes, 2C TCR rotates ~20° perpendicular to the peptide axis, resulting in alternative contacts between the two complexes. These two examples demonstrate that a TCR can recognize different peptides with minimal changes to the CDR loop conformations.

A third mechanism of TCR cross-reactivity relies upon CDR loop conformational changes in order to adjust to peptide differences. The prototypical representation of TCR cross-reactivity is the case of A6 TCR with two HLA-A2 peptides (Ding et al., 1999). A6 binds to the HTLV-1 (Human T-lymphotropic virus 1) peptide, Tax, but also cross-reacts with a yeast peptide, Telp. Comparison between A6-Tax and A6-Telp complexes showed that A6 TCR docks on the MHC in a very similar orientation. The differences between the complexes lie in the CDR3β loop where it adopts different conformations and "tunes" to different peptide ligands (Baker et al., 2012). In addition, A6 in complex with Telp partly unravels the MHC-I α2 helix conformation upon ligation, supporting a model of conformational melding between MHC, TCR, and peptide (Figure I-3). Similar to the A6 TCR complexes, in BM3.3 pMHC-TCR complexes, BM3.3 TCR also displays CDR loop conformational melding. BM3.3 was crystallized with the K^b allele and several allopeptides (pBM1, pBM8), as well as a viral peptide from vesicular stomatitis virus, VSV8 (Mazza et al., 2007a). BM3.3 CDR3α adopts large conformational changes to accommodate peptide differences between the three structures. When comparing

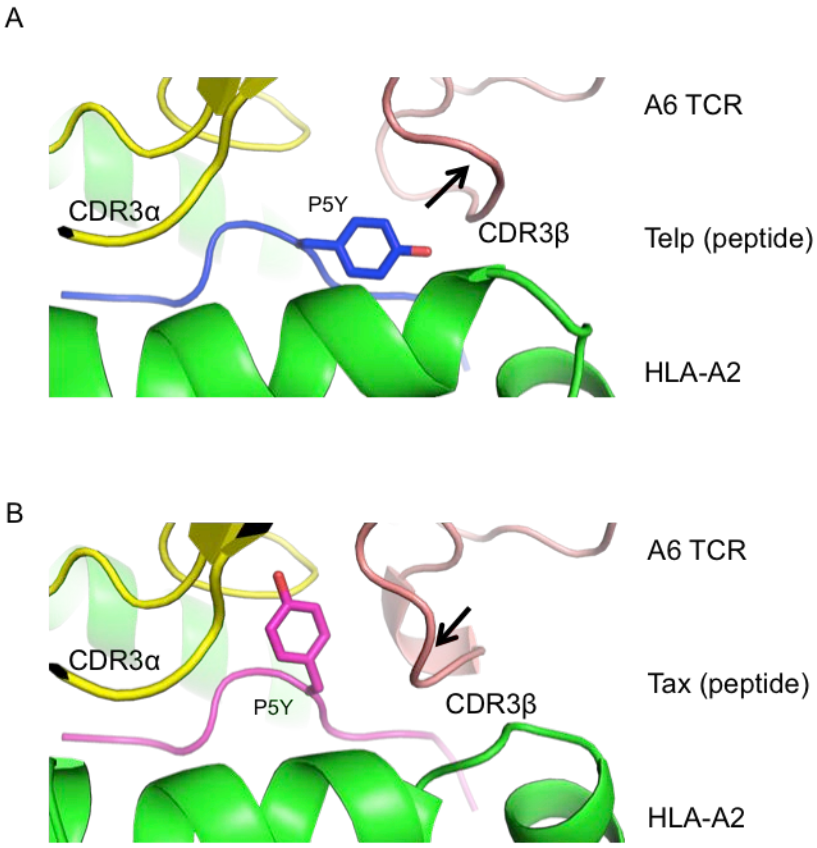

Figure I-3 Conformational Change of CDR3β Loop Between Two Ligands.
(A) The same TCR, A6, bound to two different peptides, Telp (blue, PDB: 3H9S) and-Tax (B) (pink, PDB: 1AO7) shows rearrangements in peptide sidechains and CDR3 β loop. MHC-α2 helix is also slightly altered after ligand binding.

pMHC complexes before and after TCR ligation, the structures show that peptide side chain conformations also change upon ligand binding (Borbulevych et al., 2009). Taken together, the same TCR can cross-react with different peptide-MHC complexes by rotations, translations, CDR3 conformational changes, and pMHC conformational changes.

It is important to note, that although CDR loop conformational changes can occur after pMHC-TCR ligation, it is not a feature of all ligand binding and is not present in all examples of pMHC-TCR structures (Rossjohn et al., 2015). Comparing the unliganded 1E6 TCR to 1E6 in complex with HLA-A2 and a preproinsulin peptide, the TCR CDR loop conformations do not change upon binding (Bulek et al., 2012).

The mechanisms described above are derived from investigating models of alloreactivity and T cell mediated autoimmunity, however, the molecular mechanisms of TCR cross-reactivity in the context of viral peptides remains to be further investigated. Heterologous immunity is the immunity that develops towards one pathogen when exposed to an unrelated pathogen (Welsh et al., 2010). In the case of the tuberculosis vaccine, bacille Calmette-Guerin (BCG) and Vaccinia virus (VacV), mice with immunity to BCG are also protected against subsequent VacV challenge (Mathurin et al., 2009). Similarly, mice immune to Lymphocytic choriomeningitis virus (LCMV) are protected against subsequent VacV challenge (Selin et al., 1998). This seminal study between

LCMV and VacV was surprising at the time because LCMV and VacV are two evolutionarily distinct viruses.

LCMV is a small ssRNA virus (~10 kilobase pairs) belonging to the *Arenaviridae* family and encodes only four transcripts: GP, NP, Z, and L. On the contrary, VacV is a large dsDNA enveloped virus (~200 kilobase pairs) belonging to the *Poxviridae* family. Subsequent studies demonstrated that cross-reactivity between these two viruses are present in CD8+ memory T cells (Cornberg et al., 2007). Cross-reactive T cell lines can be generated for further investigation *in vitro* by peptide stimulation of splenocytes from LCMV-immune mice with VacV peptide epitopes. This tool was used to identify specific cross-reactive epitopes such as LCMV GP34 and VacV A11R, both of which bind to K^b (Cornberg et al., 2010).

Of particular interest is how the same TCR recognizes two disparate epitopes. Chapter IV of this dissertation will describe our approach to identify cross-reactive TCR motifs. Furthermore, Chapter V will explore the molecular mechanisms of how an isolated TCR recognizes GP34 and A11R.

CHAPTER II

SUSCEPTIBILITY TO HLA-DM PROTEIN IS DETERMINED BY A DYNAMIC CONFORMATION OF MAJOR HISTOCOMPATIBILITY COMPLEX CLASS II MOLECULE BOUND WITH PEPTIDE

Preface

The work presented here highlights my contribution to the following published article:

Yin, L., **Trenh, P**., Guce, A., Wieczorek, M., Lange, S., Sticht, J., Jiang, W., Bylsma, M., Mellins, E.D., Freund, C., Stern, L.J. (2014). Susceptibility to HLA-DM protein is determined by a dynamic conformation of major histocompatibility complex class II molecule bound with peptide. J. Biol. Chem. 289, 23449–23464.

Author Contributions

Liusong Yin conducted peptide-binding experiments and screened for crystallization conditions of DR1-A2L9. I contributed to this work by solving the crystal structure of this complex.

Introduction

HLA-DM (DM) mediates the exchange of peptides loaded onto MHC-II molecules during antigen presentation in endosomal compartments. The differential effect of DM on different peptides has been shown to play a key role in antigen presentation (Amria et al., 2008; Lazarski et al., 2006; Lovitch et al., 2003) and epitope selection (Hartman et al., 2010; Kremer et al., 2012; Lazarski et al., 2005; Pu et al., 2002; Sant et al., 2005; Yin et al., 2012). The determinants of DM-mediated peptide exchange remain elusive and controversial (Anders et al., 2011; Ferrante and Gorski, 2010; Narayan et al., 2009; Painter et al., 2011; Pos et al., 2012; Stratikos et al., 2004).

In contrast to MHC-I molecules, peptides bound to MHC-II are extended at both N and C-termini. For this reason, peptide-binding pockets are confusingly labeled starting with N-terminal pocket 1, which does not correlate with the position along the peptide. Previous studies proposed that DM targets the conserved hydrogen bonds between MHC-II and peptide in the vicinity of the P1 pocket (Narayan et al., 2007; Stratikos et al., 2004), although others found that these interactions are dispensable for DM action (Ferrante and Gorski, 2010; Painter et al., 2011; Schulze et al., 2013; Zhou et al., 2009). Peptide side chain pocket interactions also have been implicated in DM susceptibility, particularly emphasizing the importance of pocket 1 near the N terminus of the bound peptide (Anders et al., 2011; Chou and Sadegh-Nasseri, 2000).

Recently, a structure of DM bound to HLA-DR1 (DR1) carrying a covalently bound, N-terminally truncated peptide showed extensive rearrangement of the MHC-II pocket 1 (Pos et al., 2012). The interpretation is that this model represents a transient intermediate in the DM-MHCII-peptide exchange reaction, which would resolve upon peptide binding and displacement of DM. A model was presented in which peptides that lack an optimum pocket 1 residue would not be able to displace DM from the DM-DR complex, tying DM-susceptibility to the nature of the residue that occupies pocket 1. In this study, we directly evaluated the role of the side chain occupying the pocket 1 in determining susceptibility to DM-mediated peptide exchange.

We focused on variants of the alloantigenic peptide HLA-A2$_{104–117}$ bound to the human class II MHC protein DR1 (HLA-DRA*01:01, DRB1*01:01). This MHC-peptide complex has been shown to be the target of graft rejection in solid organ transplant recipients (Hanvesakul et al., 2007; Smith et al., 2011) and allospecific human monoclonal antibodies (Wölpl et al., 1998). HLA-A2$_{104–117}$ is the predominant endogenous peptide bound to DR1 isolated from a B lymphoid cell line (Chicz et al., 1992), presumably reflecting its high binding affinity, resistance to DM-mediated exchange, and the abundance of the HLA-A2 protein in the endosomal subcellular compartment(s), where peptide binding and DM editing occur. The crystal structure of DR1-HLA-A2$_{104–117}$ has been solved previously defining the major and minor anchor residues (Figure II-1) (Murthy and Stern, 1997), which makes it an ideal example to study DM action.

Results

To address the role of pocket 1 in DM susceptibility, we designed variant peptides with non-optimal P1 pocket side-chains and combined these with optimal residues at other pockets. We reasoned that if we were able to find a pocket 4, pocket 6, or pocket 9 substitution that restored the wild-type (WT) affinity to a pocket 1-substituted variant, comparison of such peptides with the WT could illuminate the role of the pocket 1 residue in DM-mediated peptide exchange, because the WT peptide gains most of its binding affinity through the pocket 1 interactions, whereas the double-substituted peptides would derive most of their affinity elsewhere.

We tested each A2 peptide variant using a fluorescence polarization inhibition assay with tight-binding indicator peptide $HA_{306–318}$ (K_D = 10 nM) (Roche and Cresswell, 1990). The WT (HLA-$A2_{104–117}$) peptide bound to DR1 very tightly with IC50 values of 36 nm (Figure II-1B and Table 1), consistent with a previous report (Chicz et al., 1992). As expected (Sato et al., 2000), single amino acid substitutions of the pocket 1 side chain generally reduced binding as indicated by increased IC50 values (Figure II-1B). We selected such four variants for continued work: L1, V1, T1, and A1. Substitution of pocket 4 and pocket 6 side chains with preferred residues did not reduce IC_{50} values, and these positions were not investigated further (Hammer et al., 1993). However, we found that leucine substitution at pocket 9 (L9) substantially enhanced the binding affinity as indicated by reduced IC50

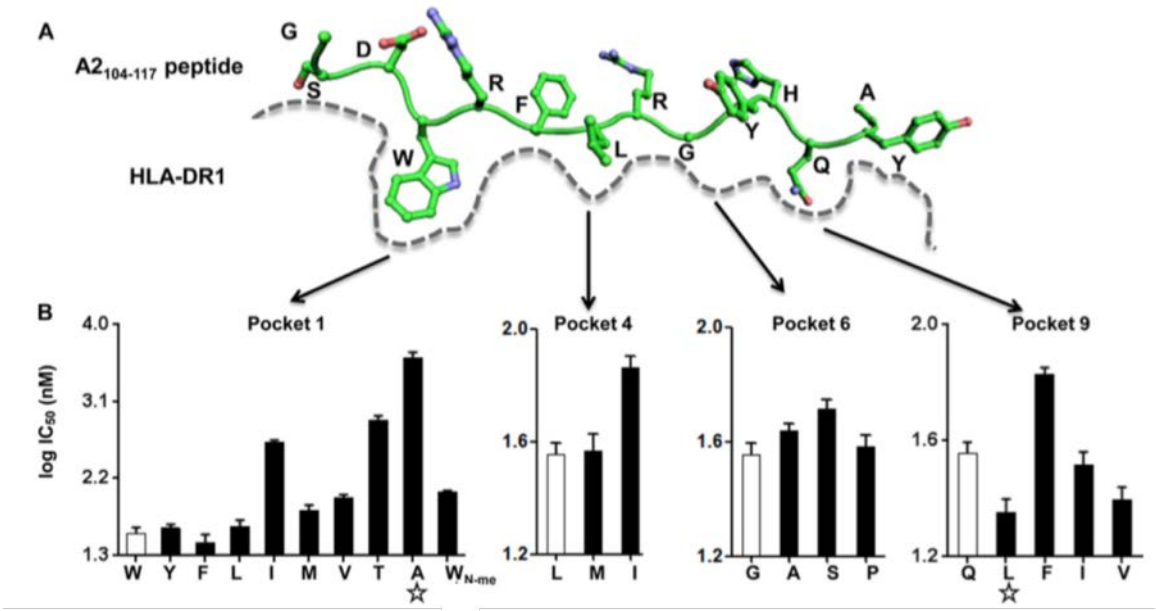

Figure II-1 Identification of peptides with weakened pocket 1 interactions and strengthened interactions at other pockets.

A) Schematic view of peptide binding groove from previously solved crystal structure of peptide HLA-$A2_{104-117}$ bound to DR1 (PDB code 1AQD). Pockets 1, 4, 6, and 9 on DR1, which harbor the major anchor residues of HLA-$A2_{104-117}$, are indicated.

B), competition binding with indicator peptide Alexa488-$HA_{306-318}$ to DR1 was measured for HLA-$A2_{104-117}$-derived peptides harboring single mutation at pockets 1, 4, 6, and 9 anchor residues. The amino acid after substitution for each peptide was labeled at the x axis, with the open bar indicating wild-type sequence, and the star indicating peptides selected for following detailed studies. These data represent three independent experiments with two replicates each.

Peptide Name [a]	Sequence [b]	IC_{50} [c] (nM)
WT	G S D **W** R F **L** R **G** Y H **Q** Y A	35.9 ± 2.2
L1	- - - **L** - - - - - - - - - - -	43.3 ± 2.9
L1L9	- - - **L** - - - - - - - **L** - -	25.2 ± 2.5
V1	- - - **V** - - - - - - - - - - -	93.5 ± 2.6
V1L9	- - - **V** - - - - - - - **L** - -	29.9 ± 3.8
T1	- - - **T** - - - - - - - - - - -	743.1 ± 21.5
T1L9	- - - **T** - - - - - - - **L** - -	46.2 ± 3.8
A1	- - - **A** - - - - - - - - - - -	3964 ± 117
A1L9	- - - **A** - - - - - - - **L** - -	36.9 ± 2.9

Table II-1 Binding and kinetic properties of HLA-DR1-peptide complexes formed with HLA-A2(104 –117)-derived peptide mutants

a Peptides are named with the substituted anchor residues.

b The pockets 1, 4, 6, and 9 anchor residues in wild type peptides (WT) and corresponding substituted anchor residues are highlighted in bold and underlined.

c 50% inhibition concentration calculated from binding competition curves.

values (Figure II-1), consistent with previous reports that leucine is an optimal residue at this position (Hammer et al., 1993; Stern et al., 1994).

The L9 substitution that enhanced binding for the wild-type peptide also enhanced binding for the pocket 1 substituted variants L1, V1, T1, and A1. In each case, substitution of leucine at position 9 into a peptide carrying a substitution at position 1 (A1L9, T1L9, V1L9, and L1L9) reversed the effect of the pocket 1 substitution and returned the peptide binding affinity to nearly the wild-type level, as judged by a competition binding assay (Table II-1). In the most extreme example, the A1 substitution showed more than 100-fold increase in IC50 relative to the wild-type peptide (Fig 1B), as expected based on the major anchoring function of the pocket 1 residue (Murthy and Stern, 1997), and consistent with a previously reported low affinity pocket 1 anchorless peptide derived from $HA_{306–318}$ (Natarajan et al., 1999).

When the A1 peptide was reconstituted with leucine at pocket 9 (A1L9), it rescued the binding affinity to a similar level as WT (Table II-1). Together, these data suggest that stronger interactions at the C-terminal end of the peptide (pocket 9) can compensate for the weaker interactions at the N-terminal end of peptide (pocket 1) in terms of binding to MHCII. One possible explanation for the unexpected role of L9 interaction could involve a potential shift to a different binding register with the A1L9 alanine not located in pocket 1. Peptide binding in a noncanonical reversed orientation has been observed, with

DM able to accelerate the exchange of $CLIP_{106–120}$ either canonically or inversely bound to DR1 (Günther et al., 2010; Schlundt et al., 2012).

To confirm that A1L9 binds to DR1 with the same binding register as the WT peptide, we determined the x-ray crystal structure of DR1-AlL9 to a resolution of 2.2 Å (Figure II-2 and Table II-; PDB ID 4OV5). The structure shows A1L9 bound to DR1 in the canonical orientation with the expected pocket 1 and pocket 9 residues (Figure II-2 A and B), with no substantial differences in peptide conformation or MHCII side chain orientation relative to WT (Figure II-2C), including residues implicated in DM action (αTrp-43, αPhe-54, βAsn-82, and βPhe-89) (Figure II-2D). As expected from previous work, weakening interactions in the pocket 1 results in a decreased MHCII-peptide lifetime and increased susceptibility to DM. However, we found that weakened P1 interactions could be completely compensated by substitutions elsewhere in the peptide; in particular, reconstitution with leucine at pocket 9 restores the conformation similar to that of DR1-WT.

Discussion

In the study where this work was published, further studies support that DM-susceptibility is correlated with DR1 conformation. We measured peptide dissociation kinetics in the absence or presence of and showed that DM accelerated the dissociation of peptides differentially. The L1, V1, T1, and A1 peptide complexes all showed an

increase in DM susceptibility (data not shown). Surprisingly, L1L9, V1L9, T1L9, and A1L9 all demonstrated DM susceptibility similar to WT. This indicates that a gain of

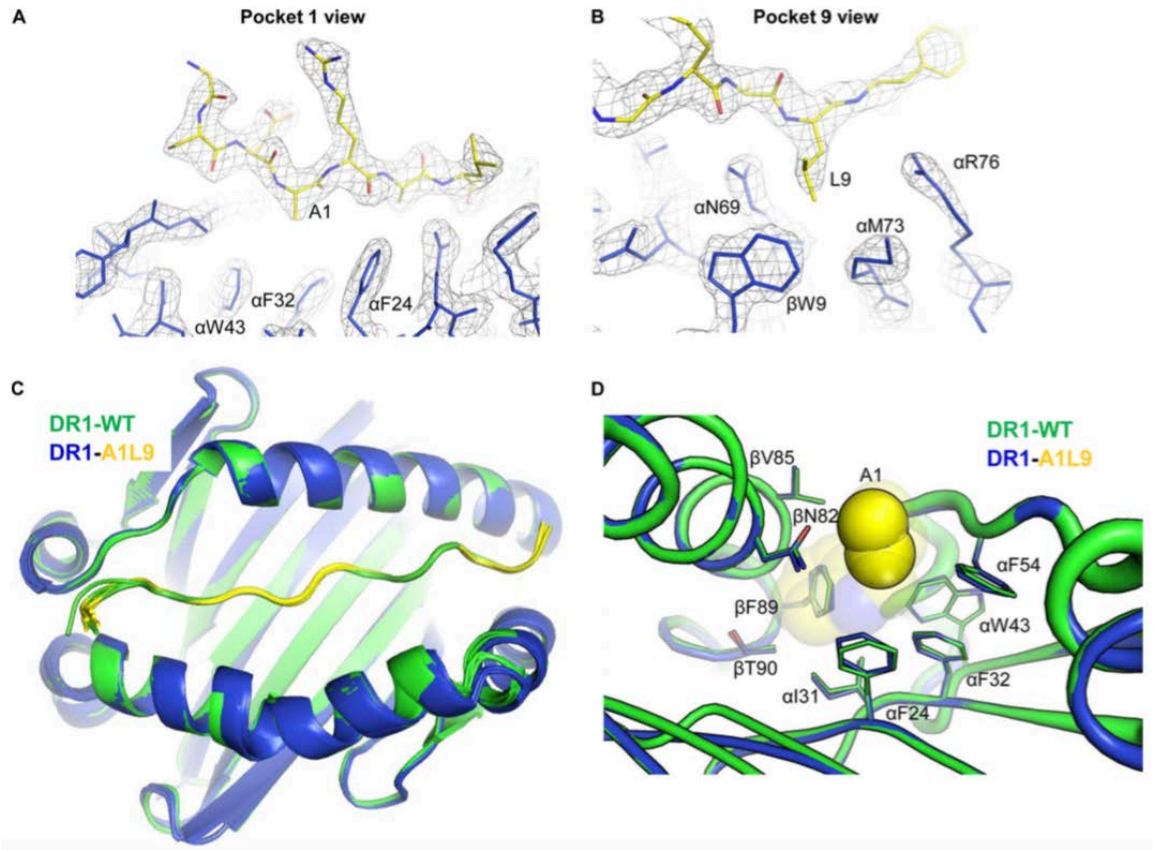

Figure II-2 X-ray crystal structure of DR1-A1L9 confirms the expected peptide - binding motif and reveals similar conformation to that of DR1-WT.
Electron density for peptide and DR1 residues around pocket 1 (*A*) and pocket 9 (*B*) with the peptide anchor residue and key contact residues on DR1 indicated in x-ray crystal structure of DR1-A1L9 complex (PDB code 4OV5). *C*, top view of DR1-A1L9 structure overlapped with DR1-WT structure. Four molecules in the asymmetric unit of DR1-WT and six molecules in the asymmetric unit of DR1-A1L9 are shown. DR1 in DR1-A1L9 is colored *blue*, A1L9 *yellow*, and DR1-WT *green*. *D*, a closer view at the pocket 1 with key residues is indicated. The pocket 1 alanine (A1) in DR1-A1L9 is highlighted as a *yellow* bulky residue, and the tryptophan in DR1-WT is highlighted as a transparent shaded residue. Statistics of x-ray data collection and refinement are listed in Table II-2.

Data collection and refinement statistics

	DR1-A1L9 (4OV5)
Data collection	
Space group	P1 21 1
Cell dimensions	
a, *b*, *c* (Å)	95.96, 173.19, 96.48
α, β, γ (°)	90.00, 109.72, 90.00
Resolution (Å)	46.00-2.20 (2.24-2.20)[a]
R_{merge} (%)	4.8 (47.7)
$I/\sigma I$	16.09 (1.44)
Completeness (%)	93.7 (81.8)
Redundancy	2.5 (2.3)
Refinement	
Resolution (Å)	46.00-2.20 (2.24-2.20)
No. of reflections	140,577
R_{work}/R_{free}	20.52/23.95
No. of atoms	19,082
Protein	18,443
Water	639
B-factors	33.25
Protein	33.30
Water	31.77
Root mean square deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.773
Ramachandran favored (%)	98.08
Allowed region	1.88
Outlier region	0.04

[a] Values in parentheses are for highest resolution shell.

Table II-2 Data Collection and Refinement Statistics of DR1-A1L9 Complex

interactions at the C-terminal end can compensate for loss of interactions at the N-terminal end of the peptide.

SDS sensitivity, antibody binding, hydrodynamic measurements, SAXS, and NMR analyses revealed conformational differences between DR1-A1 and DR1-A1L9. Judged by several criteria, the DR1-A1 complex with weakened pocket 1 interaction appears to adopt a conformation different from conventional DR1-peptide complexes, which is more susceptible to DM-mediated peptide exchange. Reconstitution with leucine at pocket 9 restores the conformation to the less susceptible form similar to that of DR1-WT.

The idea that the pocket 1 region is crucial to DM susceptibility has been given additional prominence by a recent structure of DM bound to DR1 with a covalently trapped truncated peptide (Pos et al., 2012), in which DR1 conformational alterations around pocket 1 were observed and associated with DM interaction. However, we have shown here that the A1 and A1L9 peptides, both with the same non-anchoring pocket 1 alanine, have dramatically different DM susceptibilities. Both of these peptides were able to efficiently displace bound peptides, in the absence or presence of DM, with the equilibrium level reached according their binding affinities and not their pocket 1 residues. Thus, DM susceptibility is not determined by interactions only in the pocket 1 region of the binding sites but rather by interactions throughout the peptide-binding groove.

These data provide the clearest evidence yet that DM edits conformations, not sequences, and helps to define a molecular basis for conformational editing. In conclusion, we have shown that the DM susceptibility of DR1-peptide complexes can be determined by peptide substitutions outside the P1 pocket and by MHC substitutions away from peptide and DM interaction sites. Our data support the idea that the dynamic conformation of MHCII-peptide complexes determined by interactions throughout the peptide binding groove is the key determinant of DM susceptibility and that DM acts as a conformational rather than sequence editor.

CHAPTER III

TCRαβ CHAIN PAIRING AND CDR3 SEQUENCES REGULATE MHC-TCR CONTACTS AND LIGAND SPECIFICITY

Preface

The work presented here highlights my contribution to the following published journal articles:

Stadinski, B.D.*, **Trenh, P**.*, Smith, R.L., Bautista, B., Huseby, P.G., Li, G., Stern, L.J., and Huseby, E.S. (2011). A role for differential variable gene pairing in creating T cell receptors specific for unique major histocompatibility ligands. Immunity 35, 694–704. *Authors contributed equally

Stadinski, B.D., **Trenh, P**., Duke, B., Huseby, P.G., Li, G., Stern, L.J., and Huseby, E.S. (2014). Effect of CDR3 sequences and distal V gene residues in regulating TCR-MHC contacts and ligand specificity. J. Immunol. 192, 6071–6082.

Author Contributions

Huseby lab conducted TCR cloning, biochemical, and cellular assays. I screened for crystal conditions of I-Ab-3K-J809.B5 complex. Brian Stadinski screened for crystallization conditions of I-Ab-3K-14.C6 complex. I contributed to this work by solving the crystal structures of both complexes.

Abstract

A limited set of T cell receptor (TCR) variable (V) gene segments are used to create a repertoire of TCRs that recognize all major histocompatibility complex (MHC) ligands within a species. The antigen specificity of TCRs is created from the combinatorial pairing of one of a set of germline encoded TCR variable (V) alpha and V beta gene segments with randomly created CDR3 sequences. How individual αβTCRs are constructed to specifically recognize a limited set of MHC ligands is unclear. Here we have identified a role for the differential paring of particular V gene segments in creating TCRs that differentially recognize MHC class II. Structural studies support biophysical experiments and indicate that TCR specificity for MHC ligands is not driven by germline encoded pairwise interactions. Rather, identical TCRβ chains can have altered peptide-MHC (pMHC) binding modes when paired with different TCRα chains. Furthermore, CDR3α residues can modulate the geometry in which TCRs bind pMHC, governing whether and how germline encoded TCR Vα and Vβ residues interact with MHC. These findings demonstrate that the specificity of individual T cell clonotypes arises not only from TCR residues which create direct contacts with the pMHC, but also from a collection of indirect effects which modulate how TCR residues are used to bind pMHC.

Introduction

Within a population, MHC allele diversity greatly limits the ability of pathogens to escape immune responses (Košmrlj et al., 2010; Messaoudi et al., 2002). However, the diversity of MHC creates a unique problem for generating host-MHC restricted mature T cell repertoires. Prior to selection, TCR rearrangement has to generate a collection of TCRs, which in aggregate have the ability to recognize any of the possible MHC classes and alleles present within a species. To mature, pre-selection thymocytes expressing these TCRs are subjected to the selective pressures of positive selection to ensure T cells are MHC restricted, and negative selection to limit autoimmunity (Fink and Bevan, 1995; Kappler et al., 1987; Kisielow and Von Boehmer, 1995; Mathis and Benoist, 2004). After selection, individual TCRs must be MHC class specific, allowing a division of labor between MHC class II-reactive helper T cells and MHC class I-reactive cytotoxic T cells (Babbitt et al., 1985; Zinkernagel and Doherty, 1974).

The sequence of individual TCR clonotypes is created from the pairing of one of a limited set of Vα and Vβ gene segments with highly variable CDR3α and CDR3β sequences derived from V(D)J recombination (Bevan, 1977). Through studying T cells isolated from mice with limited negative selection, we have provided evidence that T cells can have a range of pMHC cross-reactivity patterns (Huseby et al., 2005a, 2006; Stadinski et al., 2011). These and other T cell activation studies suggest that TCRs may have an intrinsic ability to bind pMHC, which is regulated by TCR V gene pairing or

CDR3 sequences (Blackman et al., 1986; Merkenschlager et al., 1997; Rubtsova et al., 2009; Stadinski et al., 2011; Stritesky et al., 2013; Zerrahn et al., 1997).

Most TCRs bind MHC ligands within a semi-conserved diagonal orientation, which largely places the CDR3 loops atop the bound peptide and the germline encoded V gene CDR1 and CDR2 residues positioned over the MHC alpha helices (Rudolph et al., 2006b). Examination of TCRs carrying similar TCR V genes engaged to similar MHC alleles have shown a more limited range of TCR-pMHC docking angles (Dai et al., 2008; Feng et al., 2007). These structural observations have led to a hypothesis that particular germline encoded residues of TCR V genes have been evolutionarily selected to bind MHC in conserved ways and provide TCRs with a built-in specificity for MHC ligands (Garcia et al., 2009b; Marrack et al., 2008).

To test this hypothesis and identify mechanisms that control TCR ligand specificity, we isolated and characterized a set of TCRs reactive to IA^b-3K carrying the identical TCRβ chain. The TCRs were either self-tolerant or self-reactive, and differed in the ability to cross-react with other classes of MHC ligands. By comparing how each receptor bound the same pMHC complex, we sought to identify how TCR chain pairing affects MHC specificity. X-ray crystallographic experiments indicated that MHC specific TCRs and MHC class cross-reactive TCRs utilizes a spectrum of pMHC binding modes within a conventional docking footprint. Our data demonstrate that differential TCR α-chain pairing can result in MHC specific TCRs that have altered TCRβ CDR loop

conformations and placements, as well as modified TCRβ-MHC contacts. By comparing similar TCRs with alternate CDR3α sequences, we demonstrate that residues distal to the binding interface can have profound affects on the overall geometry, thereby indirectly affecting how different TCR Vα and TCR Vβ residues interact with pMHC residues.

Results

We created a model to identify mechanisms that control TCR specificity by generating mice expressing the Vβ8.2 TCRβ chain of YAe62 TCR as a transgene (YAe62β mice). YAe62 is a MHC-cross-reactive TCR isolated from single peptide mice (Huseby et al., 2005b). T cells that were reactive to the 3K-peptide bound to the IA^b were isolated for further studies. To examine the influence of TCRα chain pairing and CDR3α sequences on TCR-MHC interactions, the IAb-3K-reactive TCRs J809.B5 and 14.C6 were studied. These two self-tolerant TCRs carry the identical TCRβ and TCR Vα2.8 sequences, differing only in the CDR3α sequence (Table III-1). In comparison to the self-reactive YAe62 TCR, these TCRs have the same TCRβ chain, but differing TCR α chains.

Structural comparison of an MHC specific and an MHC cross-reactive TCR

By comparing how these TCRs engage pMHC, we sought to identify how αβTCR chain pairing controls ligand specificity. To identify in detail how paring a Vα2 TCR α-chain with the YAe62 β-chain creates pMHC specificity, we first determined the 2.7Å crystal structure of the J809.B5 TCR bound to IA^b-3K and compared it to the 3.2Å structure of

TCR	Vα	CDR1α	CDR2α	CDR3α	Vβ	CDR1β	CDR2β	CDR3β
YAe62	4.12	STTGYPT	QVTTANNK	CAANSGTYQRFG	8.2	TNNHNN	YSYGAGSTEKGD	CASGDFWGDTLYFG
J809.B5	2.8	ENSAFDY	AIRSVSDK	CAASKGADRLTFG	8.2	TNNHNN	YSYGAGSTEKGD	CASGDFWGDTLYFG
14.C6	2.8	ENSAFDY	AIRSVSDK	CAASRDSGQKLVFG	8.2	TNNHNN	YSYGAGSTEKGD	CASGDFWGDTLYFG

Table III-1 TCR Sequences of Receptors Used for Crystallographic Studies

the YAe62 TCR bound to IA^b-3K (Table III-2) (Dai et al., 2008). The shared use of the identical Vβ chain provides a direct way to evaluate how different Vα pairings control ligand specificity. The overall orientation and size of the interfaces of the two complexes were highly similar (Figure III-1A-E) Additionally, the conformation of the TCR Cα, and TCR Cβ domains, and the TCR Vα and TCR Vβ framework regions, superpose closely. The total buried surface area (BSA) for the J809.B5 complex is approximately 1560 $Å^2$, comparable to 1410 $Å^2$ for the parent YAe62 complex. In both cases the TCRβ chain contributes the majority of contacts to IA^b-3K, 74% of the total BSA for the J809.B5 TCR, versus 67% for YAe62 TCR (Figure III-1F-G). Some conformational differences were observed for the TCRα subunit, which was expected due to the alternate α-chain usage (Figure III-2). Thus, as was observed with the energetic footprint analysis (data not shown), the J809.B5 TCR does not create self-tolerance from gross structural alterations.

Alternative TCRα to pMHC contacts by the J809.B5 TCR versus the YAe62 TCR explains the change in peptide recognition of J809.B5. The majority of the YAe62 TCRα contacts with IA^b-3K involve CDR1α Y29 engaging the IA^bβ chain residues T77 and H81 (Figure III-2B, blue). This MHC-dominant binding mode allows the YAe62 TCR to be less-sensitive to alanine substitutions at several residues of the peptide. In contrast, the J809.B5 TCRα does not make any substantial contacts with the IA^bβ chain (Figure III-2, orange), consistent with the energetic binding data (data not shown).

The J809.B5 TCR binds IA^b-3K with an altered TCRβ-pMHC binding mode

Data Collection	
Space group	C 1 2 1
Cell dimensions	
a, b, c (Å)	239.855, 73.016, 65.669
α, β, γ (°)	90.00, 90.707, 90.00
Resolution (Å)	30-2.70 (2.75-2.71)
Rmerge (%)	8.4 (54.0)
I/σI	18.6 (2.8)
Completeness (%)	98.7 (97.6)
Redundancy	4.0 (4.0)
Refinement	
Resolution (Å)	30-2.70
No. reflections	31144
Rwork/Rfree (%)	21.81/25.95
No. atoms	
Protein	6120
B-factors	
Protein	51.5
Rmsds	
Bond lengths (Å)	0.008
Bond angles (°)	1.140
Ramachandran favored (%)	95.6
Allowed region	4.0
Outlier Region	0.4%

*Highest resolution shell is given in parenthesis

Table III-2 Data Refinement and Statistics of 809.B5 TCR in Complex with I-A^b-3K

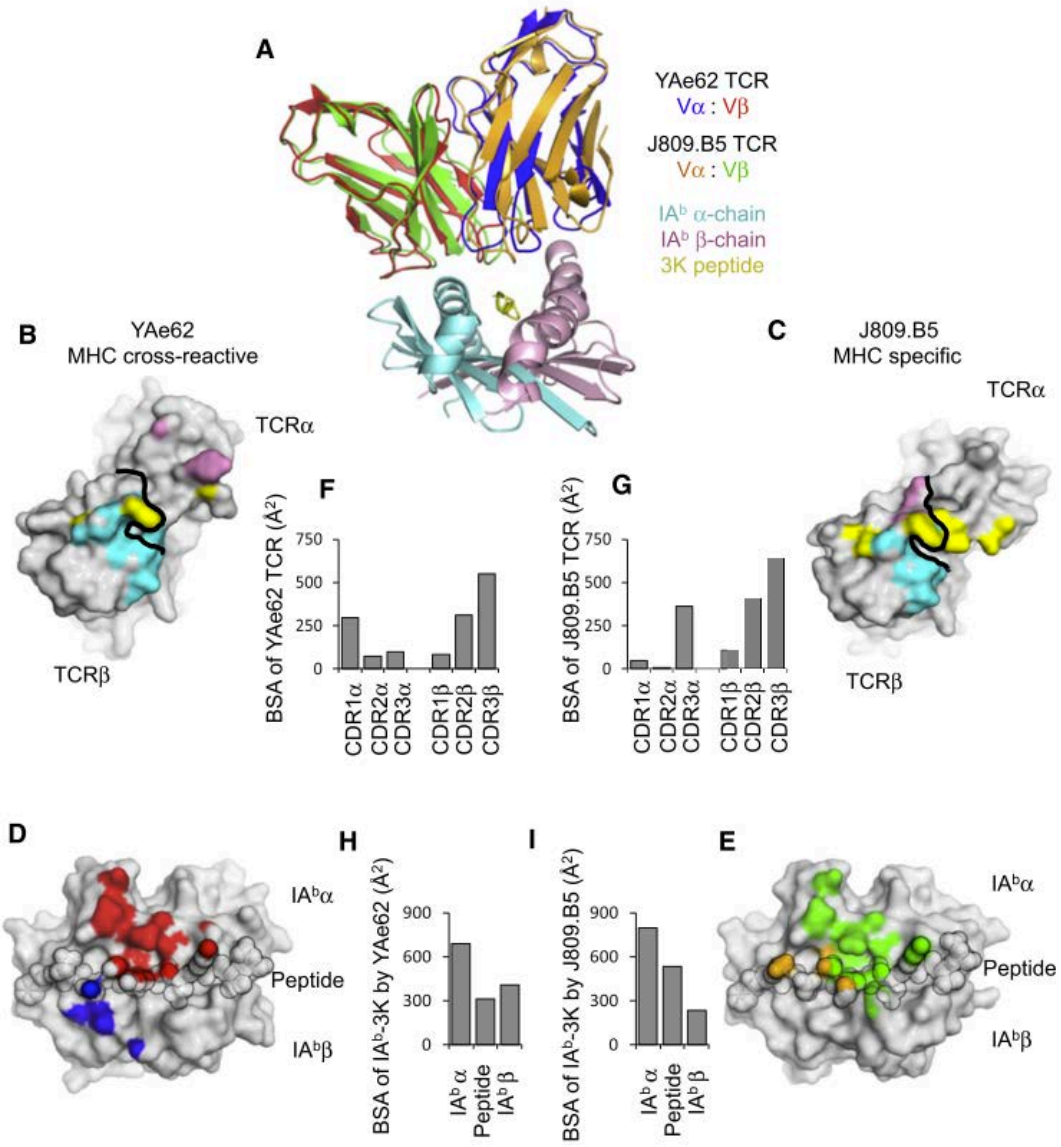

Figure III-1 MHC cross-reactive and MHC specific TCRs bind IAb-3K within a similar footprint.

(A) Overlay of YAe62 and J809.B5 TCRs binding IAb-3K. The YAe62 TCR is colored red (TCRβ) and blue (TCRα); the J809.B5 TCR is colored green (TCRβ) and orange (TCRα). IAb-3K is colored cyan (IAbα chain), yellow (peptide) and magenta (IAbβ chain).
(B, C) Projection of IAb-3K binding onto the (B) YAe62 TCR, or (C) J809.B5 TCR. Contacts with the IAbα chain (colored cyan), peptide (colored yellow) and the IAbβ chain (colored magenta). Black line demarcates border of the TCRα subunit from the TCRβ subunit.

Figure III-1 MHC cross-reactive and MHC specific TCRs bind IAb-3K within a similar footprint (cont.)

(D, E) Projection of the (D) YAe62 TCR or (E) J809.B5 TCR binding onto IAb-3K. YAe62 TCRα contacts are colored blue, YAe62 TCRβ contacts are colored red. The J809.B5 TCRα contacts are colored orange and the TCRβ contacts are colored yellow. The peptide residues are outlined in black.
(F–I) The amount of Buried Surface Area (BSA) contributed by the (F) YAe62 or (G) J809.B5 TCRα and TCRβ loops to the binding reaction with IAb-3K. (H) The amount of BSA contributed by the peptide or MHC chains for the binding reaction with YAe62 TCR or the (I) J809.B5 TCR. Figures were made using PyMol (DeLano).

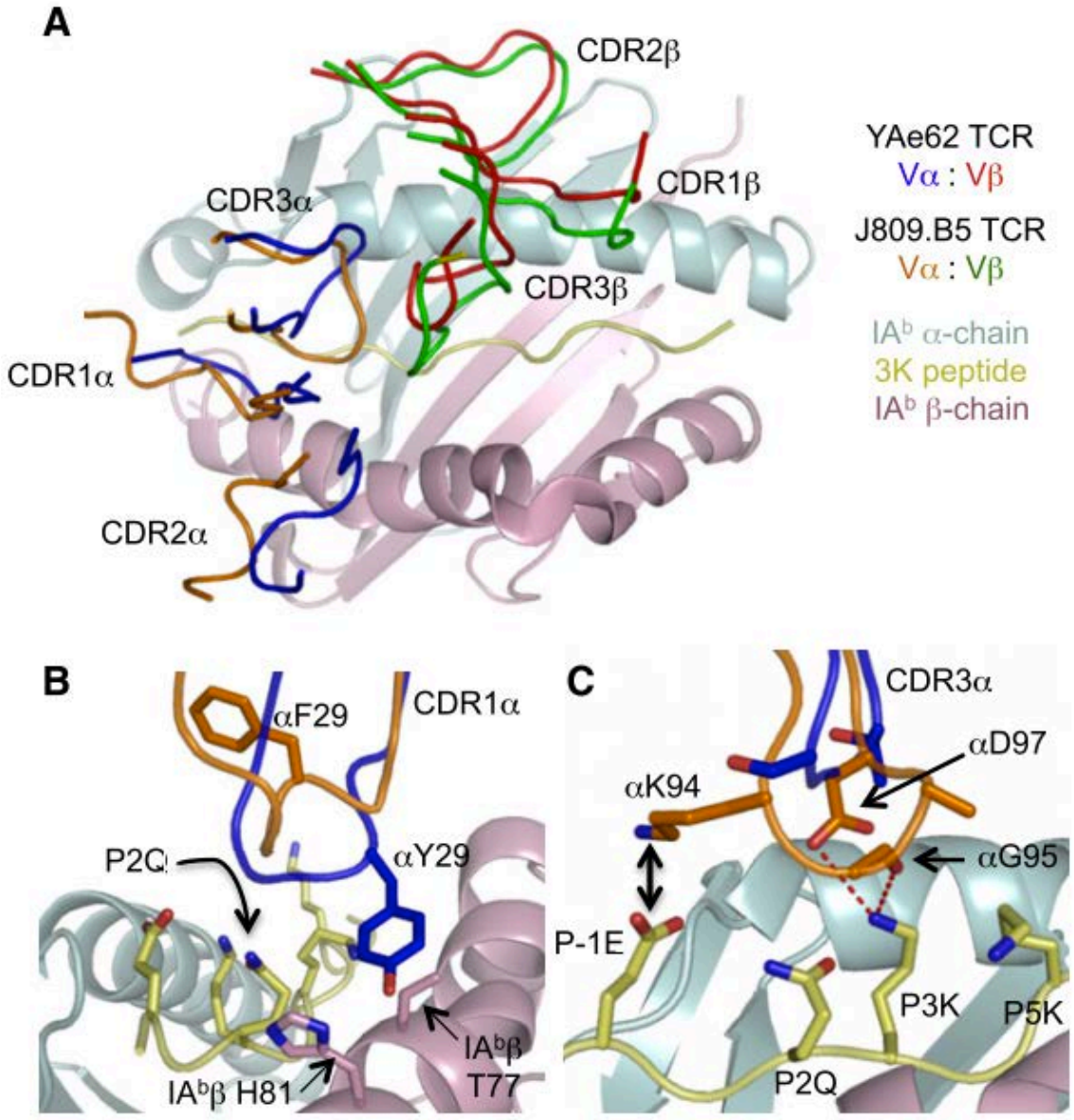

Figure III-2 The J809.B5 TCRα chain induces a TCR rigid body movement and replaces the YAe62 TCRα contacts to MHC with TCRα contacts to the peptide.
(A) A rigid body movement rotates the J809.B5 TCRβ chain 1–2 Å towards the peptide; J809.B5 TCRβ (green), J809.B5 TCRα (orange), YAe62β (red) and YAe62α (blue).
(B) The YAe62 CDR1 αY29 (blue) contacts IAb βT77 and βH81 whereas as the J809.B5 CDR1α (orange) makes no contact with the IAbβ chain.
(C) The J809.B5 CDR3α (orange) makes extensive contacts with the P-1E and P3K residues of the peptide whereas the YAe62 CDR3α (blue) makes very few. Salt bridge (black arrow) and hydrogen bonds (red dashed lines) between J809.B5 CDR3α and the 3K peptide are shown. Figures were made using PyMol (DeLano).

Despite having the identical TCRβ sequence, the TCRβ chains within the J809.B5 and YAe62 TCRs bind IA^b-3K differently. The J809.B5 Vα2 chain pairing results in the TCRβ binding IA^b-3K with a modified CDR3β loop conformation, and a rigid body movement that rotates the TCRβ chain 1–2 Å towards the center of the interface (Figure III-2), compare J809.B5 CDR3β (green) with YAe62 CDR3β (red)). In conjunction with the modified TCRβ-pMHC binding mode a set of TCRα to TCRβ inter-chain contacts are observed between TCR CDR1α, CDR2α and CDR3α residue side chains (orange) with the CDR3β loop (green) (Figure III-3). Most prominent are contacts between CDR1α Y31 and CDR2α R50 with CDR3 Jβ D97, and between CDR3 Jα A96 with CDR3 Jβ W95, contacts that are not present in the MHC cross-reactive YAe62 TCR. The J809.B5 TCR and YAe62 TCR structural comparison demonstrates that one control mechanism which impacts ligand specificity is the ability to modify how TCR CDR loop residues bind pMHC.

<u>CDR3α sequences can modulate CDR1 and CDR2 interactions with pMHC</u>

To gain insight into how CDR3α sequences can influence CDR1 and CDR2 interactions with MHC, we determined the 3.2Å cocrystal structure of the 14.C6 TCR bound to IAb-3K, and compared it to J809.B5 bound to the same ligand. The J809.B5 TCR and the 14.C6 TCR engage IA^b-3K similarly but with a marked change in orientation, such that there is a 7.5° difference in the tilt of the TCR relative to pMHC (Figure III-4, Table III-3). This change in TCR-pMHC binding geometry is confirmed using an R-free analysis of the TCR V domains. The 14.C6:IA^b-3K complex has an R-free of 23%

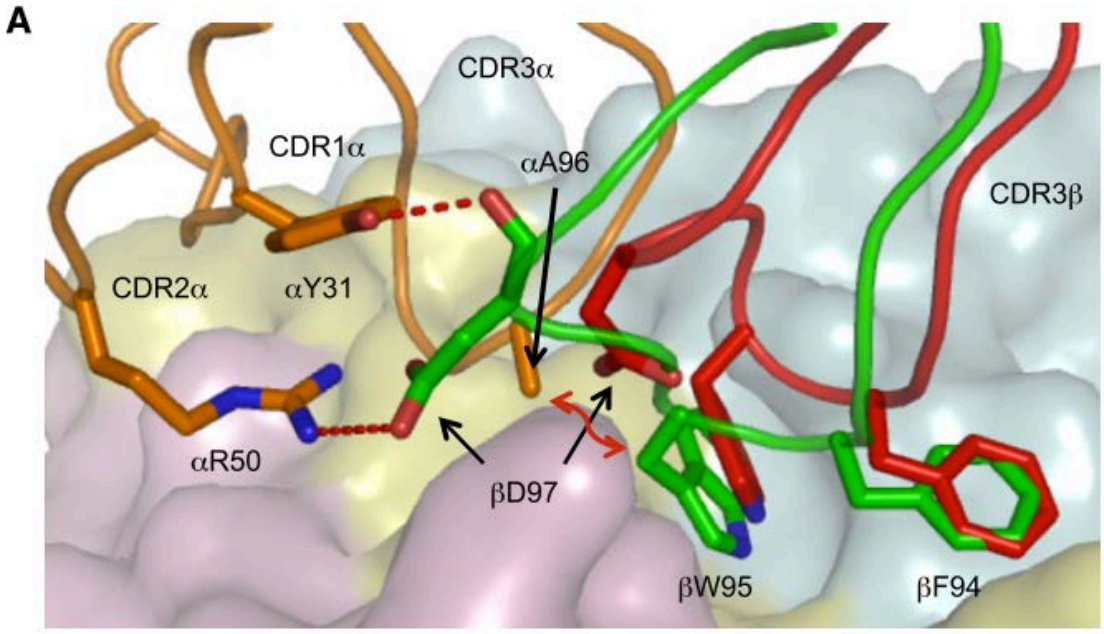

Figure III-3 TCRα chains can induce a conformational change to the TCRβ CDR3 loop.
(A) The J809.B5 chain (green) is in an altered conformation as compared to the identical CDR3β chain within the YAe62 TCR (red). The J809.B5 CDR3β chain creates a hydrogen bond (red dashed lines) with J809.B5 CDR1 αY31, CDR2 αR50 and a Van der Waals contact (red arrow) with CDR3 JαA96.

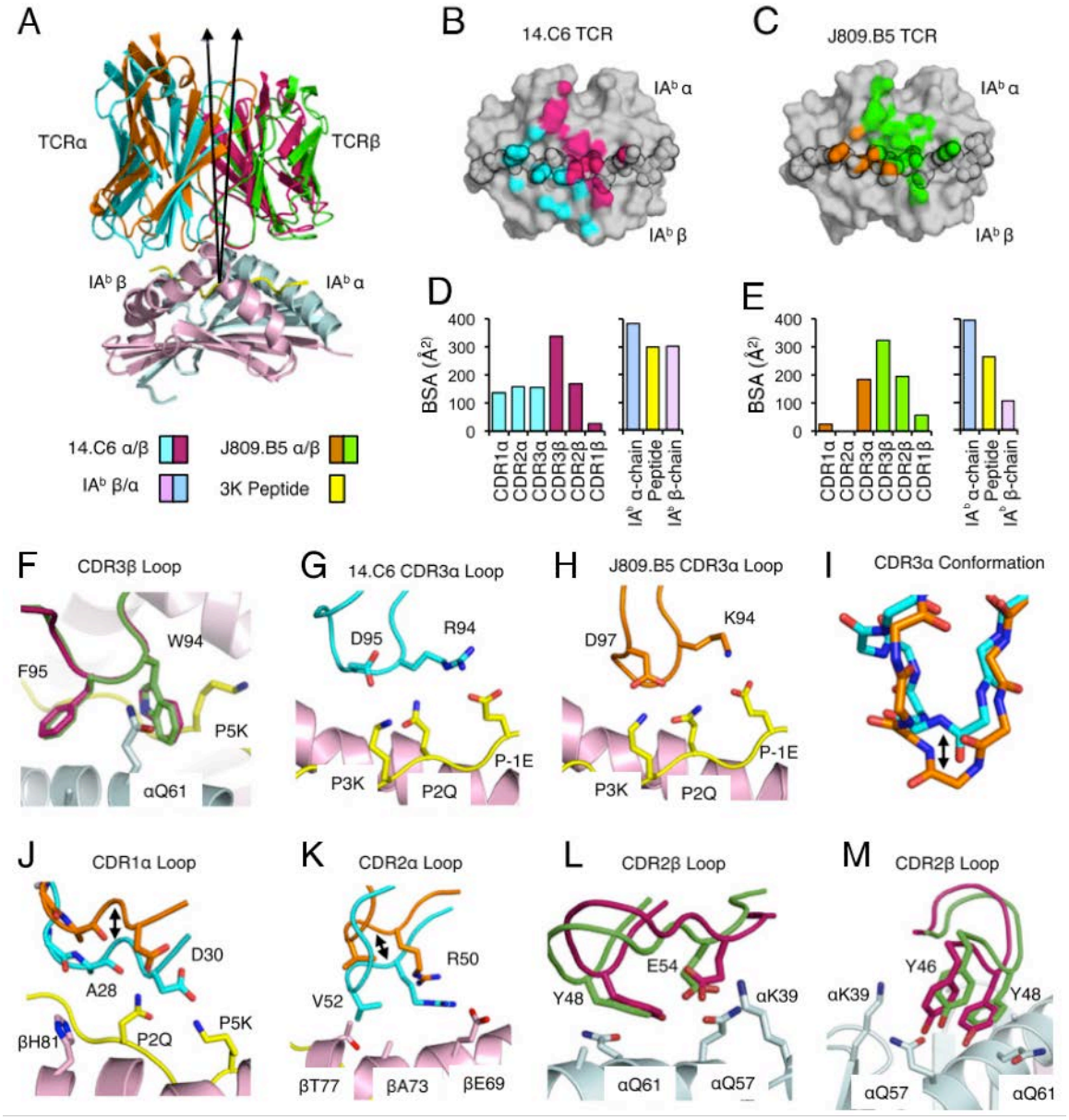

Figure III-4 CDR3α Sequence Can Alter the Positioning of CDR1 and CDR2 Loop Atop MHC.

(**A**) Overlay of the 14.C6 and J809.B5 TCRs binding IA^b-3K. The 14.C6 TCR is colored red (TCRβ) and blue (TCRα); the J809.B5 TCR is colored green (TCRβ) and orange (TCRα). IA^b-3K is colored cyan (IA^{b}α chain), yellow (peptide), and magenta (IA^{b}β chain). A 7.5% difference in the tilt of the TCR docking geometry atop IA^b-3K is observed based on a vector from the center of mass of each TCR to the α-carbon of the p5 peptide residue. (**B**) Projection of the 14.C6 TCR or (**C**) J809.B5 TCR binding onto IA^b-3K. 14.C6 TCRα contacts are colored blue, 14.C6 TCRβ contacts are colored red.

The J809.B5 TCRα contacts are colored orange and the TCRβ contacts are colored green. The peptide residues are outlined in black. (**D**) The amount of buried surface area (BSA) contributed by the 14.C6 or (**E**) J809.B5 TCRα, TCRβ, peptide or MHC chains for the binding reaction with IA^b-3K. (**F**) The 14.C6 CDR3β loop (magenta) and the J809.B5 CDR3β loop (green) are in a similar conformation and make similar contacts with IA^b-3K. (**G**) The 14.C6 CDR3α loop residues R94 and D95 interact with the peptide residues, P-1E, P2Q and P3K. (**H**) The J809.B5 CDR3α loop residues K94 and D97 interact with the peptide residues, P-1E, P2Q and P3K. (**I**) The 14.C6 CDR3α loop (blue) and the J809.B5 CDR3α loop (orange) are in a different conformation when bound to IA^b-3K. (**J**) 14.C6 CDR1α residues A28 and D30 (blue) make extensive contacts with the P2Q and P5K residues of the peptide, and the IA^bβ chain residue H81. The J809.B5 CDR1α loops residues A28 and D30 (orange) are shifted approximately 4 Å, allowing only the A28 residue to make minimal contacts with peptide, and no contact with MHC. (**K**) The 14.C6 TCR CDR2α residues R50 and V52 (blue) interact with IA^b β-chain residues βE69, βA73 and βT77, contacts that are not present in the J809.B5:IA^b-3K structure (orange) due to a 4 Å rigid body shift. (**L, M**) The main-chain of CDR2β loop at the βY48 position in the 14.C6 structure (green) is shifted approximately 1Å as compared to the J809.B5 structure (magenta), and some rotamer differences are observed for the CDR2β residues E54 and Y46. Figures were made with PyMol

whereas if the 14.C6 V domain if forced into the J809.B5 docking geometry, the resulting complex would have an R-free of 42%. This change in tilt allows the 14.C6 TCR to form an increased number of contacts with the IA^bβ-chain, as compared to the J809.B5 TCR (Figure III-4B, C). In addition, the total buried surface area (BSA) for the 14.C6 complex (approximately 1970 $Å^2$) was greater than that for the J809.B5 complex (approximately 1540 $Å^2$), primarily stemming from changes to the BSA of the CDR1α and CDR2α loops and the IA^b β-chain (Figure III-4D, E). In contrast to these changes, CDR3β loop conformations and the placement of CDR3 βW94 and βF95 residues within these structures are similar (Figure III-4F).

To evaluate how CDR3α sequences can affect TCR interactions with pMHC, we compared the 14.C6 and J809.B5 CDR3α loop conformations when bound to IA^b-3K. In both structures the 14.C6 and J809.B5 CDR3α loops make similar types of contacts with the peptide (Figure III-4 G, H). Despite the 14.C6 CDR3α loop being one amino acid longer, the J809.B5 CDR3α loop extends 2.5Å further towards the peptide, relative to the 14.C6 CDR3α loop (Figure III-4I). Because of the different CDR3α loop conformations, a steric clash would occur between the J809.B5 CDR3α loop and the peptide if this TCR bound IA^b-3K

	14.C6:IA^{b}3K
Data collection	
Space group	P1 2_1 1
Cell dimensions	
a, *b*, *c* (Å)	66.007, 74.142, 259.289
a, b, g (°)	90.00, 92.00, 90.00
Resolution (Å)	50–3.25(3.31–3.25)
R_{merge}	15.5(53)
I / sI	12.9(3.8)
Completeness (%)	99.1(97.9)
Redundancy	6.6(3.8)
Refinement	
Resolution (Å)	47–3.26(3.38–3.26)
No. of reflections	39080 (3819)
R_{work} / R_{free}	18.63/23.62
No. of atoms	
Protein	12366
Water	
B-factors	
Protein	65.4
Water	
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.666
Ramachandran favored (%)	95.55
Allowed region	4.13
Outlier region	0.32

Table III-3 Data Collection and Structure Statistics of 14.C6 TCR in Complex with I-A^b-3K

with the same geometry as that of the 14.C6 TCR. Thus, the differences in CDR3α loop conformation likely underlie why the J809.B5 and 14.C6 TCRs bind IA^b-3K with different tilts. The most prominent differences in how the J809.B5 and 14.C6 TCRα chains engage IA^b-3K involve CDR1α and CDR2α residues. The 14.C6 TCR docking geometry allows its CDR1α loop to make extensive contacts with the P2 and P5 positions of the 3K peptide, as well as the IA^bβ chain residue βH81 (Figure III-4J). In contrast, the J809.B5 CDR1α loop is positioned such that it makes only minimal contacts with the 3K peptide and no contacts with MHC (Figure III-4J). Likewise, the 14.C6 TCR CDR2α residues interact with IA^bβ chain, whereas these contacts are absent in the J809.B5:IA^b-3K structure due to the 4 Å rigid body shift (Figure III-4K).

Difference in the requirements of TCR Vβ and IA^bα residues for J809.B5 and 14.C6 T cells to undergo activation, arise from less dramatic structural changes. The main-chain of the J809.B5 and 14.C6 CDR2β loops are shifted approximately 1Å at the CDR2 βY48 position. In addition, some rotamer differences are observed for the CDR2β residues, βY46 and βE54; while the IA^b αQ57 and αQ61 side chain rotamers are similarly orientated (Figure III-4L, M). Thus, the requirements of TCR and MHC side chains for T cell activation can be strongly affected by modest shifts in CDR loop placement and side chain rotamer differences.

Discussion

The identical set of TCR V genes are used to create TCRs that are specific for either classical or non-classical MHC class I or MHC class II ligands. Although there are reports of some TCR V genes being preferentially used in either $CD4^+$ or $CD8^+$ T cell populations (Jameson et al., 1990; Sim et al., 1996), none of the V genes are precluded from creating MHC class I or MHC class II specific receptors (Garman et al., 1986; Jorgensen et al., 1992; Valkenburg et al., 2010). Thus, TCRs are able to utilize a limited set of TCR V gene residues to create receptors that bind the highly diverse set of ligands. The dichotomy that individual TCR chains can be incorporated into receptors that are able to bind a wide range of polymorphic MHC ligands, while complete receptors can specifically engage a limited set of MHC ligands, strongly suggests αβTCR chain pairings limit MHC cross-reactivity. The isolation of MHC specific and MHC cross-reactive, IA^b-3K reactive TCRs carrying the identical TCRβ chain allowed us to identify mechanisms by which TCRs become MHC specific.

Several mechanisms likely allow TCRαβ chain pairing to create receptors that are specific for MHC ligands. These could include the binding requirements of each TCR chain for the pMHC, modification of CDR loop conformational structures, CDR loop positioning or induced alterations of the TCR or pMHC complexes that arise during binding (Jorgensen et al., 1992). Not surprisingly, the self-tolerant, MHC specific TCRs carrying the YAe62β chain required binding energy to be contributed from multiple residues of the peptide at likely TCRα contacts. However, MHC specific TCRs did not

require additional MHC contacts at potential TCRα binding sites. Within the J809.B5 TCR-pMHC structure, no contacts were observed between the CDR1α and CDR2α loops with IA^b-3K. Our data further demonstrated that TCRα pairings, which created MHC specific TCRs, had a modified TCRβ binding reaction with pMHC. Thus, TCRα chains did not influence pMHC specificity by requiring additional MHC class or allele specific contacts. This was a surprising observation given the strong Vα2 gene segregation to MHC specific TCR, and suggested this skewing was not based on TCR Vα gene residues providing contacts with the pMHC.

Unlike immunoglobulin genes, TCRs do not undergo somatic hypermutation. Therefore, the only way TCRs can be created with reactivity for antigens presented by host-MHC alleles is through the combinatorial pairing of different TCR V gene segments and the creation of random CDR3 sequences through V(D)J recombination. The major force in creating TCRs with diverse antigen reactivity patterns occurs from V(D)J recombination, as this process determines both the length of the CDR3 loops and the identity of the amino acids present within the central portion of the pMHC binding site. Sequence differences amongst the Vα and Vβ gene segments also contribute to creating TCRs with a diversity of ligand specificities. Differential TCR V gene pairing creates TCRs that express different amino acid sequences at potential sites of pMHC contact. In addition, CDR1 and CDR2 loops can adopt one of a set of unique canonical structures, which allow different TCR Vα and Vβ pairings to create TCRs with a modest range of tertiary structures (Al-Lazikani et al., 2000; Arden, 1998; Garcia et al., 1999).

For the J809.B5 and 14.C6 TCRs structures bound to IA^b-3K studied here, CDR3α sequences are shown to regulate the extent in which the CDR1α and CDR2α loops contribute to the binding reaction. The range of TCR sequence-dependent changes in TCR V gene-MHC interactions are mirrored by observations that the MHC allele and the bound peptide or lipid sequence can influence how TCR CDR1 and CDR2 loops engage MHC residues (Adams et al., 2011; Borbulevych et al., 2011a; Deng et al., 2012; Felix et al., 2007; Liu et al., 2013; López-Sagaseta et al., 2012; Mazza et al., 2007b; Reiser et al., 2003; Stadinski et al., 2011). These findings may explain how very modest changes in TCR sequence, including TCRs with only a single amino acid change in the CDR3α loop, can distinguish TCRs expressed on $CD4^+$ T cells recognizing MHC-II ligands from TCRs on $CD8^+$ T cells recognizing MHC-I ligands (Correia-Neves et al., 2001). Collectively, the ability of CDR3 sequences and peptide sequences to modulate how TCR V gene residues engage MHC residues provides a mechanism whereby the TCR-MHC binding properties of the mature T cell repertoire can be tailored during T cell selection to recognize antigens displayed by host-MHC molecules (Berg et al., 1989; Bevan, 1977; Morris et al., 2011; Scott et al., 1989).

CDR1 and CDR2 residues can make direct contacts with the pMHC ligand. In addition, CDR residues that do not contact the pMHC can affect the ability of TCRs to adopt or stabilize different pMHC binding conformations. The indirect effects created by TCR inter- and intra-chain interactions that regulate antigen specificity from positions outside

of the binding interface can be seen as analogous to the allosteric regulation of protein function, that is controlled by V(D)J recombination and TCR V gene pairing (Nussinov and Tsai, 2013). The common occurrence of interactions between CDR loops suggests that this is a general feature of TCRs, and that one of the functions of combinatorial diversity and T cell development is to equip mature T cells with TCRs that stabilize CDR loop conformations that specifically recognize antigens presented by host MHC.

CHAPTER IV

TCR Vβ AND CDR3 SIGNATURES OF HETEROLOGOUS IMMUNITY BETWEEN LCMV AND VACV

Author Contributions

Keith Daniels infected mice with LCMV and harvested splenocytes for further analysis.

Abstract

Memory T cells are reactivated upon secondary encounter with cognate ligand, but also can be reactivated by recognition of heterologous epitopes. An example is the heterologous CD8 T cell response that occurs between Lymphocytic choriomeningitis virus (LCMV) and Vaccinia virus (VacV). Many studies of cross-reactive T cell recognition involving different viral epitopes or self-epitopes typically focus on particular individual T cell receptors, and so the numbers and diversity of T cell receptors that participate in heterologous immune responses *in vivo* are generally unknown. Increasingly, next generation sequencing (NGS) is used to characterize TCR repertoires, but it is unclear how the data can be used to identify cross-reactive TCRs. Here, we analyzed the full breadth of the $A11R_{198-206}$-cross-reactive component of the LCMV-immune TCRβ repertoire directly *ex-vivo* by tetramer staining and deep sequencing. We identified a highly diverse population of many hundreds of distinct TCRβ clonotypes able to recognize both LCMV GP_{34-41} and VacV $A11R_{198-206}$. Within this population we observed several distinct sequence motifs with signatures that include both Vβ gene and CDR3 sequence components. Different signatures were observed for different CDR3 lengths, suggesting different recognition modes. Evidence was obtained for cross-reactive TCRβ populations recognizing VacV $A11R_{198-206}$and other LCMV epitopes in addition to GP_{34-41}.

Introduction

TCRs are a key component in the adaptive immune system. In contrast to conserved receptors in the innate immune system, TCRs represent a diverse set of cellular receptors from random gene segment pairings known as V(D)J recombination (Arden et al., 1995). TCR sequences are largely unpredictable because of random addition and deletion of N nucleotides between each gene segment. The TCR repertoire should be diverse enough to bind a plethora of foreign peptides, yet specific enough to avoid self-reactivity. However, TCRs are generally characterized as degenerate for their pervasiveness in ligand specificity (Borbulevych et al., 2009; Colf et al., 2007; Collins and Riddle, 2008; Jones et al., 2008; Mazza et al., 2007a; Sewell, 2012; Wucherpfennig et al., 2007). Detailed information about TCR repertoires may be informative in providing insight into ligand specificity. While a TCR can tolerate many changes to a peptide, it is unclear how to identify TCRs with overlapping specificity.

Cross-reactivity between viral epitopes may lead to heterologous immunity, which is the immune response towards a pathogen that results from previous exposure to an unrelated pathogen (Welsh et al., 2010). One example where protective heterologous immunity has been demonstrated is between two evolutionary distinct viruses: LCMV and VacV. Mice with immunity to LCMV by a previous infection are protected against subsequent VacV infection (Selin et al., 1994, 1998). Further studies identified a VacV epitope, A11R, which is uniquely expanded from LCMV infection (Cornberg et al., 2007, 2010). A11R-reactive cells are not detectable in naïve mice, suggesting that A11R-reactive cells only

arise due to cross-reactivity with LCMV epitopes (Cornberg et al., 2007; Shen et al., 2013).

One identified A11R cross-reactive epitope from LCMV is GP34, which shares only three identical residues. The use of two tetramers directed against GP34 and A11R demonstrate that the same T cells indeed recognize these two disparate epitopes. Mice that were infected with LCMV and analyzed 8 days later by a panel of antibodies for TCR Vβ genes show that the hierarchy of Vβ gene usage varies for the recognition of A11R and GP34 epitopes (Shen et al., 2013). T cells specific for GP34 are predominantly TRBV15 and TRBV13-3/13-2 (Vβ12 and Vβ 8.1/8.2 in the Arden nomenclature), while A11R specific cells are predominantly TRBV13-3/13-2 and TRBV12-2/12-1 (Vβ 5.1/5.2). Interestingly, when splenocytes from LCMV-immune animals were expanded with A11R peptide, these cell lines were dominated by TRBV15 gene usage (Cornberg et al., 2007, 2010). Furthermore, when a TRBV15-dominant A11R cross-reactive cell line was adoptively transferred into naïve mice, mice were protected from VacV challenge. Because TRBV is a consistently identified from different studies, these data suggest that overlapping GP34 and A11R TRBV15 T cells may represent a biologically important population.

Of interest is a detailed characterization of the GP34/A11R repertoire to elucidate specific cross-reactive determinants. With standard Sanger sequencing methods, it is often difficult to capture the breadth of a TCR repertoire. The advent of NGS has

allowed many groups to identify and characterize specific TCR subtypes. Availability and accessibility of these tools has fostered complex analysis that extends our understanding of the immune system (Laydon et al., 2015; Rosati et al., 2017). However, it is unclear how to isolate and identify T cells from cross-reactive populations. Here, we present a tetramer-based approach followed by NGS analysis. We have identified various dissimilar cross-reactive signatures between the GP34/A11R epitopes that are only apparent after parsing our data by TCR Vβ gene and CDR3 length.

Results

Approach to Identify Cross-reactive Receptors in LCMV-immune Mice with Tetramer Sort and NGS

To identify GP34 and A11R cross-reactive TCRs directly *ex vivo*, we isolated splenocytes from LCMV-immune animals and used an NGS approach to identify and quantify TCRβ CDR3 sequences (Figure IV-1). Splenocytes from 15 mice were pooled because GP34- and A11R-reactive cells comprise a small percentage of the memory population (~4% and ~0.4% respectively). To isolate specific TCRs, pooled splenocytes were separated using single tetramer stains: GP34, A11R, and a control tetramer. We chose this method instead of a double tetramer or bispecific molecule (Shen et al., 2010) because our single tetramer stains could be reliably separated above the background and could be used without affinity and avidity competition. Live cells were stained with CD8+, CD44+, and tetramer reagents, sorted by FACS, followed by subsequent genomic

DNA extraction. Tetramer negative cells were also sorted and sequenced to compare the extent of heterogeneity between different populations and samples.

To identify cross-reactive clonotypes, we separated clonotypes that were present in both GP34 and A11R tetramer-positive populations separately rather than evaluating clonotypes that were mutally exclusive. Non-crossreactive clonotypes are sequences that were only observed by one tetramer and not the other. The dominant Vβ genes used in GP34+ are TRBV15-01, TRBV16-01, and TRBV13-02, whereas the dominant Vβ in A11R+ is TRBV03-01 (Figure IV-1B, C). This data supports and extends previous studies using antibody staining to characterize T cell populations. For example, for GP34-reactive cells, our data is supported by TRBV15-01 and TRBV13-02 gene usage observed. However, A11R-reactive cells are dominated by TRBV03-01 clonotypes, a Vβ gene that was not tested previously. The distribution of GP34/A11R cross-reactive cells from each population matches the respective distribution prior to separating these clonotypes by cross-reactivity, indicating that cross-reactivity is not limited to particular Vβ genes.

We were interested in comparing the relative populations of GP34 and A11R-reactive T cells to establish that these populations are distinct from each other. Many tools available for analysis regard each unique DNA sequence as a unique clonotype. However, we were interested in functional clonotypes: TCRs that are identical at the amino acid level. From our dataset, we noticed that if we indexed our data by CDR3 sequence instead of

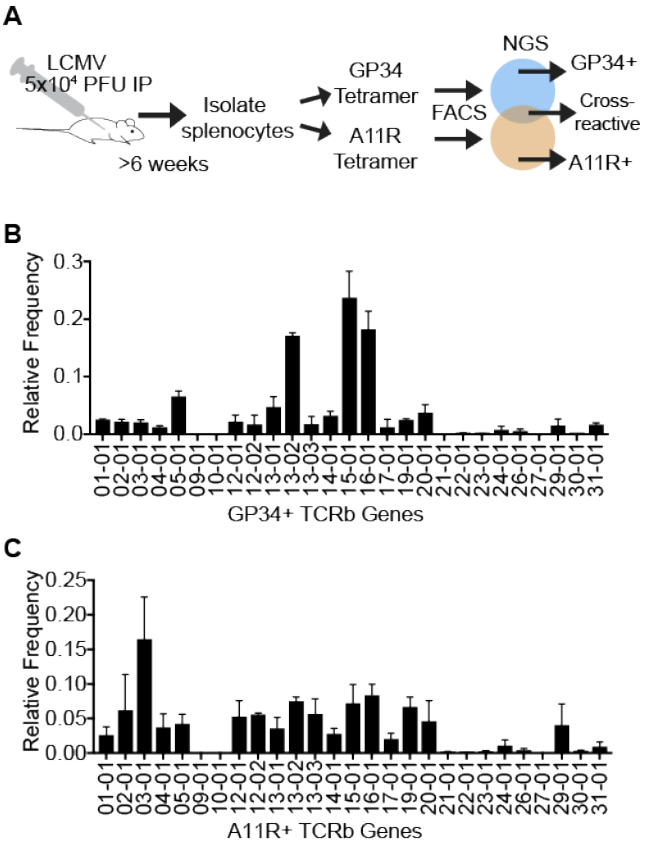

Figure IV-1 – Approach to Isolate Cross-reactive GP34/A11R T cells from LCMV-Immune Mice

(A) 15 mice were injected with a sublethal dose of LCMV I.P. After mice cleared infection and established immunity, splenocytes were isolated and pooled for analysis *ex vivo*. Samples were divided for tetramer staining, followed by sorting by FACS. Genomic DNA was extracted and analyzed for TCR sequencing by NGS. TCR sequeneces that were identified in both pools were considered cross-reactive. (B, C) TCR β gene usage for GP34 or A11R reactive cells show distinct distributions.

nucleotide sequence, the same CDR3 sequence is encoded by several DNA sequences. Additionally, the same CDR3 amino acid sequence was present in different Vβ families, which have different CDR1 and CDR2 sequences. To analyze functional clonotypes, we reindexed our data so that TCRs with the same V gene and CDR3 amino acid sequence were indexed as the same clonotype and the numbers of templates read were summed together.

First, we established that tetramer+ cells were distinct from tetramer- cells. Tetramer- populations from the GP34- and A11R- have a strong correlation; high abundance clones in GP34- are also highly abundant in A11R- (R^2 = 0.89) (Figure IV-2A). When comparing GP34 or A11R tetramer+ against tetramer- samples, tetramer+ samples were distinctly different from tetramer- samples, indicating that we separated and sequenced distinct populations with either GP34 or A11R tetramers (R^2 correlation coefficient of 0.001 and 0.353 respectively). The population of GP34+ vs A11R+ tetramer clonotypes is spread out along the diagonal (R^2 = 0.07B) than GP34- vs A11R- tetramer clonotypes. The weak association between GP34+ and A11R+ tetramer populations indicated that each tetramer stained different populations, despite originating from the same pool of mice in this study.

Only a few CDR3 sequences have been previously identified for the GP34/A11R cross-reactive T cell response. Among them are two TRBV15-01 clonotypes: A22 TCR, for which we determined the crystal structure (Chapter V), and L/a11-4, a protective cell line

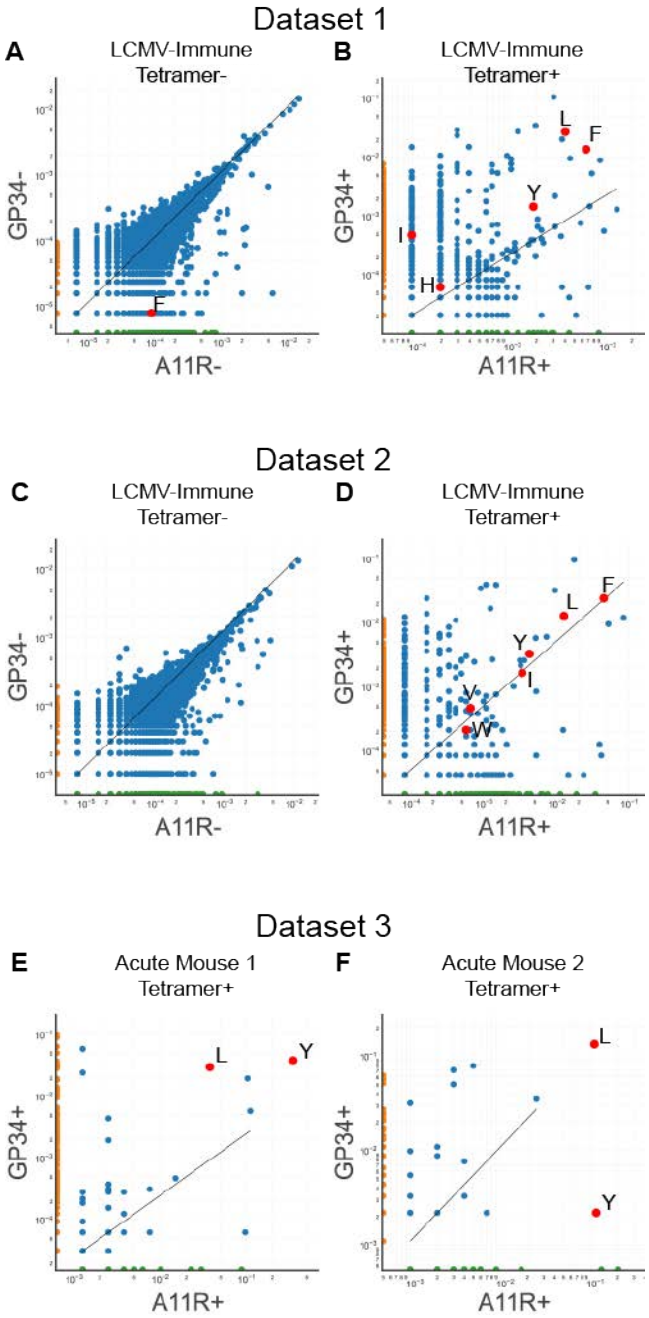

Figure IV-2 GP34 and A11R Tetramers Separate Distinct Populations
(A) GP34 tetramer negative and A11R tetramer negative populations correlate while (B) GP34 tetramer positive and A11R tetramer positive populations do not. Each spot represents a unique clonotype by Vβ gene and CDR3 sequence. The position of each spot represents the relative abundance of that clonotype as separated by respective tetramer. Data is representative of 15 LCMV-immune mice that were pooled, tetramer sorted, and TCR sequences analyzed by NGS. Red dots highlight variants of a family of receptors (A22 family) that differs by one amino acid, which is highlighted next to each

spot. Y=x line is drawn in black. (C, D) Represents a second data set, which is similar to A and B. (E, F) Two individual mice that were acutely infected are represented here by GP34 tetramer positive and A11R tetramer positive.

against VacV challenge when adoptively transferred into mice. These two clonotypes vary by one amino acid at position 4. We used the GLYPH software to identify similar clonotypes and found a family of similar receptors that differ by a single amino acid at position 5: CASSYRGGAETLY (A22 TCR), CASSLRGGAETLY (L/a11r-4), CASSFRGGAETLY, CASSHRGGAETLY, and CASSIRGGAETLY (Glanville et al., 2017). To account for possible technical variability, we repeated the NGS analysis on a separate pool of 15 LCMV-immune mice (Table IV-1). We found the A22 family consistently present in the cross-reactive population at similar relative frequencies and at low frequencies or absent in tetramer- populations (Figure IV-2).

Clonotypes belonging to the A22 family are abundant in memory T cells from LCMV-immune animals. To determine if these clonotypes are also abundant in acute LCMV-infected animals, we repeated our tetramer staining and NGS analysis on two mice acutely infected with LCMV (Figure IV-2E, F). Both mice detected high copy numbers of CASSYRGGAETLY by A11R+ tetramer, but mouse #2 had only a few copies detected by GP34+ tetramer. We have noted mouse-to-mouse variability in prior studies (Cornberg et al., 2010; Shen et al., 2013), which led to the expansion of cell lines that have differential cross-reactive epitopes. Interestingly, the CASSFRGGAETLY variant is consistently the most abundant variant detected by GP34 or A11R tetramer in both LCMV-immune datasets. Here, we note that mouse-to-mouse variability and private specificity are detectable by NGS. Taken together, variants of the A22 TCR family are represented in both acutely infected and LCMV-immune animals. The A22 family of

LCMV-Immune Dataset 1

Tetramer:	GP34-	GP34+	A11R-	A11R+
# cells	1,000,000	129,802	1,000,000	28,370
# DNA sequence reads	178,600	69,134	206,182	14,243
# DNA sequence reads (in frame)	128,290	49,116	147,280	10,220
# unique DNA sequences (in frame)	43,425	6,572	46,260	6,368
# unique Amino Acid sequences (Vβ + CDR3)	40,300	6,114	42,591	6,031
# functional clonotypes shared (tet+ & tet-) (% of total)	11941 (30%)	1132 (19%)	11941 (28%)	1132 (19%)
# functional clonotypes (only tet+ or only tet-) (% of total)	28359 (70%)	4982 (81%)	30650 (72%)	4899 (81%)

LCMV-Immune Dataset 2

Tetramer:	GP34-	GP34+	A11R-	A11R+
# cells	1000000	106222	1000000	29868
# DNA sequence reads	138140	30835	182757	15649
# DNA sequence reads (in frame)	99814	23805	132009	12225
# unique DNA sequences (in frame)	37260	1893	44629	3501
# unique Amino Acid sequences (Vβ + CDR3)	34503	1642	41074	3239
# functional clonotypes shared (tet+ & tet-) (% of total)	11109 (32%)	360 (22%)	11109 (27%)	360 (11%)
# functional clonotypes (only tet+ or only tet-) (% of total)	23394 (68%)	1282 (78%)	29965 (73%)	2879 (89%)

Acute Mouse 1

Tetramer:	GP34-	GP34+	A11R-	A11R+
# cells	100000	669025	100000	3106
# DNA sequence reads	49265	41111	19454	1083
# DNA sequence reads (in frame)	35854	32355	14272	796
# unique DNA sequences (in frame)	5027	896	2069	73
# unique Amino Acid sequences (Vβ + CDR3)	4798	816	2023	70
# functional clonotypes shared (tet+ & tet-) (% of total)	1441 (30%)	42 (5%)	1441 (71%)	42 (60%)
# functional clonotypes (only tet+ or only tet-) (% of total)	3357 (70%)	774 (95%)	582 (29%)	28 (40%)

Acute Mouse 2

Tetramer:	GP34-	GP34+	A11R-	A11R+
# cells	100000	73662	100000	4063
# DNA sequence reads	23770	1213	38831	1339
# DNA sequence reads (in frame)	17126	924	27258	1000
# unique DNA sequences (in frame)	1951	65	4047	249
# unique Amino Acid sequences (Vβ + CDR3)	1908	61	3879	236
# functional clonotypes shared (tet+ & tet-) (% of total)	1280 (67%)	19 (31%)	1280 (33%)	19 (8%)
# functional clonotypes (only tet+ or only tet-) (% of total)	628 (33%)	42 (69%)	2599 (67%)	217 (92%)

Table IV-1 Summary Table of Sequence Reads and Number of Clonotypes for Each Dataset of LCMV infected Groups

receptors is identified by a unique sequence signature and was previously shown to offer protective heterologous immune responses against VV challenge.

Cross-reactive Sequence Signatures Vary by V Gene and CDR3 Length

Because GP34/A11R cross-reactive cells are not restricted to TRBV15-01, we were interested in identifying other GP34/A11R cross-reactive clonotypes. By starting with genomic DNA instead of mRNA for NGS, one sequence read is related to the number of cells isolated. We weighed clonotypes by their respective copy numbers: a single clonotype observed hundreds of times over another are weighted so. Input files were generated for each CDR3 length for input into WebLogo software (Crooks et al., 2004). Separate inputs were generated based on GP34 or A11R tetramer staining frequency. Surprisingly, important sequence signatures were not immediately apparent.

As noted previously, CDR3 sequences may be identical in different Vβ families, so we parsed our data by Vβ gene. When clonotypes were not weighted for their abundance frequency, sequence signatures were not apparent. (Figure IV-3). However, when our data was parsed for Vβ gene, CDR3 length, and weighted for sequence reads, then cross-reactive sequence signatures became much more apparent in comparison to non-crossreactive clonotypes.

To systematically identify conserved cross-reactive sequence signatures between GP34 and A11R, we analyzed sequences separated based on the CDR3 length and Vβ using

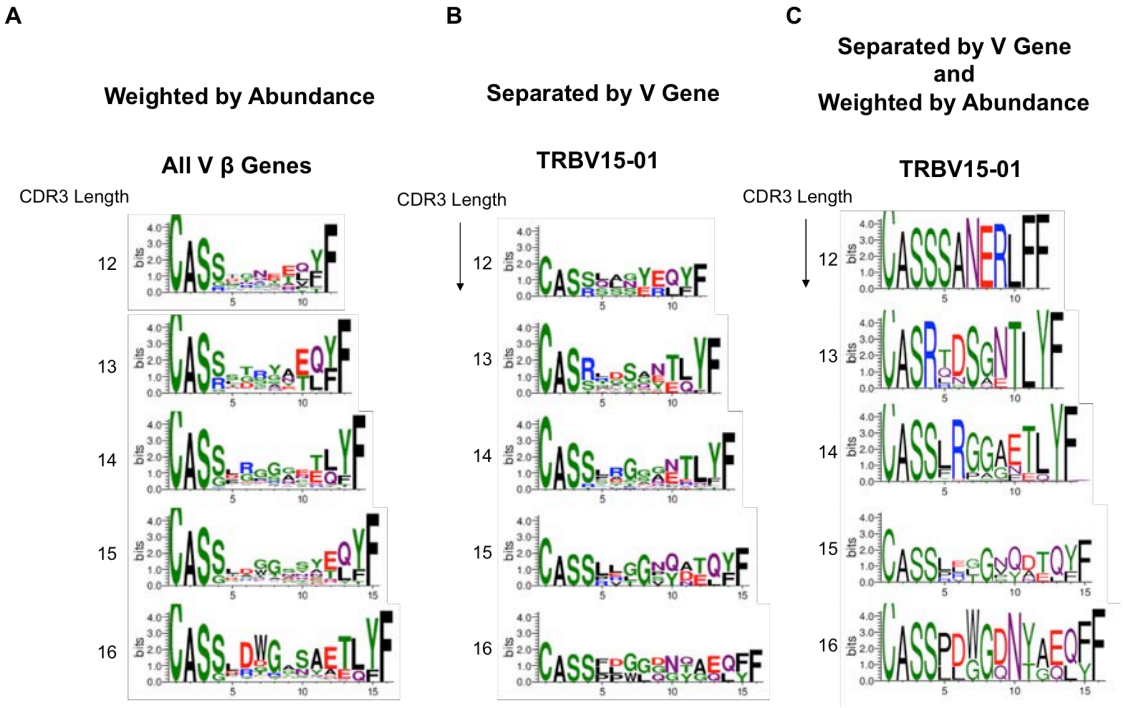

Figure IV-3 Cross-reactive Motifs are Only Apparent When Sequences are Parsed by Vβ Gene and Weighted by The Number of Sequence Reads

Sequences that were identified in both GP34 and A11R tetramer+ had different frequencies when sorted by GP34 or A11R. Representative analyses are shown here. (A) Cross-reactive sequences weighted by the number of sequence reads for each clonotype from GP34 tetramer. Sequences are separated by CDR3 length and analyzed by WebLogo, but do not show strong cross-reactive motifs. (B) Cross-reactive sequences that were separated by Vβ gene, but not weighted by sequence reads do not show strong cross-reactive motifs. TRBV15-01 shown here. (C) Cross-reactive sequences separated by Vβ gene (TRBV15-01 shown) and weighted for the number of sequence reads (GP34 tetramer shown) show unique motifs for each CDR3 length.

Two Sample Logo (Vacic et al., 2006). Two Sample Logo is based on the WebLogo algorithm but highlights significant amino acid differences between two populations. Here, we analyzed significant differences between cross-reactive (top) and non-crossreactive (bottom) clonotypes. We analyzed the top three most abundant TCR Vβs in GP34 and A11R and inferred that sequence signatures apparent in either the GP34-reactive population or A11R cross-reactive population are important for GP34-A11R cross-reactive responses. Signatures present only in the GP34-reactive populations or only in the A11R cross-reactive populations were not considered major GP34/A11R cross-reactive signatures. Interestingly, cross-reactive sequence signatures varied based on V gene and CDR3 length (Figure IV-4).

Here, we highlight sequence signatures that share three consecutive residues, enriched in both GP34+ and A11R+ cross-reactive clonotypes, and at the same positions, while allowing for one gap (Figure IV-5). In TRBV15-01, the most abundant TCR Vβ detected by the GP34 tetramer, showed two unique signatures: $R_4T_5D_6S_7X_8N_9T_{10}L_{11}Y_{12}$ and $R_6(P/G)_7G_8A_9(E/F)_{10}T_{11}L_{12}Y_{13}$ at CDR3 lengths 13 and 14 amino acids respectively. TRBV03-01, the most abundant TCR Vβ detected by the A11R tetramer, had a signature of $(L/F)_5D_6R_7A_8(N/Q)_9$ at length 14 amino acids. For the other TCR Vβ genes analyzed, TRBV13-02's unique signature is at length 14 amino acids ($D_5R_6D_7R_8X_9G_{10}$), TRBV16-01 at length 13 and 14 amino acids ($L_5G_6F_7S_8Y_9$), ($P_7I_8Y_9A_{10}$) respectively, and TRBV19-01 at length 16 amino acids ($Y_{11}A_{12}E_{13}$). Four signatures had CDR3 lengths of

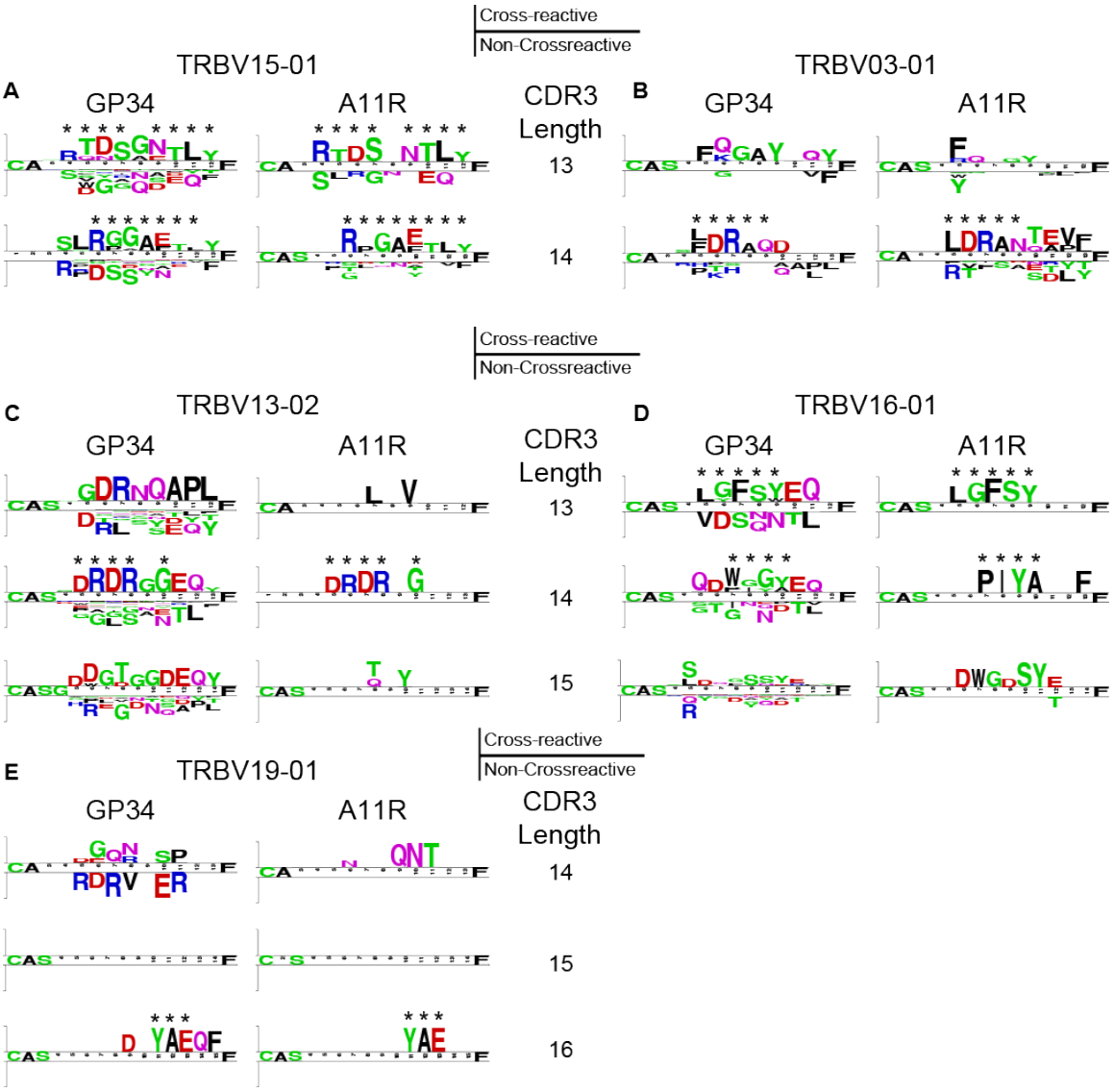

Figure IV-4 Two Sample Logo Analysis Shows Unique Cross-reactive Signatures That Are Different Between CDR3 Lengths and Different Between TCRβ Genes
Sequences were separated by CDR3 length and Vβ gene. Each unique clonotype was weighted by the number of sequence reads for that clonotype based on the number of sequence reads separated by respective tetramers (GP34 or A11R). Clonotypes that were identified by both GP34 and A11R tetramer were separated as cross-reactive (top) and clonotypes that were unique to respective tetramer sorted populations were separated as non-crossreactive (bottom). Two Sample Logo analysis shown where residues are each position that are significantly enriched (t-test with Bonferroni correction) are shown. The top three most abundant Vβ genes used in GP34 or A11R reactivity are shown: (A) TRBV15-01, (B) TRBV03-01, (C) TRBV13-02, (D) TRBV16-01, and (E) TRBV19-01. Positions with significantly enriched residues in cross-reactive populations that are present in both GP34 and A11R tetramer sorts are highlighted with a star.

		Position										
V gene	CDR3 length	3	4	5	6	7	8	9	10	11	12	13
TRBV03-01	14			L/F	D	R	A	N/Q				
TRBV13-02	14			D	R	D	R		G			
TRBV15-01	14				R	P/G	G	A	E/F	T	L	Y
TRBV15-01	13		R	T	D	S		N	T	L	Y	
TRBV16-01	13			L	G	F	S	Y				
TRBV16-01	14					P	I	Y	A			
TRBV19-01	16									Y	A	E

Figure IV-5 Summary Table of Cross-reactive Signatures Significant in GP34/A11R Cross-reactivity

14 amino acids, two of 13 amino acids, and one of 16 amino acids. Each TCR Vβ gene analyzed presented a unique cross-reactive signature that was only detectable after separating our sequences by CDR3 length, indicating that both CDR3 length and TCR Vβ genes play a role in TCR specificity and cross-reactivity.

Discussion

A TCR repertoire can generate an immense about of diversity from V(D)J recombination, yet TCRs are often considered degenerate and plastic for its ability to recognize multiple ligands. Comprehensive studies have typically focused on using a peptide-centric approach with large peptide libraries that focus on a few TCRs. In this study, we used a TCR-centric approach to characterize cross-reactivity between two epitopes. This approach is unique because only a few cross-reactive epitopes have been identified *ex-vivo*. Reliably identifying cross-reactive T cells remains a challenge due to low affinity binding to reagents and low copy numbers of cells in a single animal.

We chose to focus on LCMV-immune animals because its repertoire may be representative of a potential recall response in heterologous immunity. Our system uniquely allows us to pool syngeneic animals to increase the number of specific splenocytes without potentially introducing bias by *ex vivo* peptide stimulation. We recognize that many systemic biases can skew repertoire analysis and many groups are logically introducing new tools to analyze TCR repertoires (Dash et al., 2017; Glanville et al., 2017; Izraelson et al., 2018; Rosati et al., 2017). Our approach takes advantage of

a quantitative pipeline (Carlson et al., 2013; Robins et al., 2010), and by using Two Sample Logo to help identify motifs, we can use statistical significance to identify specific residues that are enriched in tetramer+ over tetramer- populations.

Several findings from this study contribute to our understanding of TCR cross-reactivity. First, sequence motifs are identifiable by parsing our data by TCR Vβ gene and CDR3 length. This informs us that germline encoded CDR1 and CDR2 sequences significantly contribute to epitope recognition and that unique sequences are generated based on CDR3 lengths. This is particularly surprising because it was observed that CDR3 residues are frequently positioned at similar positions in MHC recognition. However, our data suggest that residues conserved in particular interactions with specific epitopes may vary based on CDR3 length. Based on the diversity of Vβ genes used and the diversity of CDR3 cross-reactive signatures, many cross-reactive recognition modes were detectable. Surprisingly, cross-reactive sequences were not significantly biased to Vβ genes or motifs.

Although GP34/A11R cross-reactive cell lines are reproducibly generated from LCMV-immune animals, a thorough repertoire analysis has never been formally demonstrated. Interestingly, A11R reactive cells are only detectable after LCMV infection. Here, ~19% of A11R reactive clonotypes are identifiable as GP34 reactive, which suggest that A11R is exclusively cross-reactive with other LCMV epitopes besides GP34. This is surprising because the GP33/34 epitope is among the immunodominant LCMV epitopes and

represents ~20% of the CD8+ T cells during an acute infection (Van Der Most et al., 2003). In LCMV-immune animals, GP34-reactive T cells represent ~4% of the splenocytes, while ~0.4% splenocytes are A11R-cross-reactive (Cornberg et al., 2010), suggesting that most or all A11R cross-reactive cells could be GP34 reactive. However, identifying other epitopes cross-reactive with A11R remains to be elucidated.

At present, our analysis considers TCRβ clonotypes and focuses on shared identified clonotypes. Inclusion of TCRα with TCRβ as a functional TCR clonotype could considerably narrow any overlaps, as long as the sequencing depth is a representative population. A previous study estimated that ~14000 VV precursors in a naïve mouse are expanded (Seedhom et al., 2009). Given that <0.4% of splenocytes respond to A11R in LCMV-immune or VV-acute mice, we would estimate ~56 A11R clonotypes. Our current study would estimate ~3000-6000 A11R clonotypes in LCMV-immune animals. While NGS is a sensitive and thorough analysis, it is not a substitution for functional studies. Instead, it is a complementary approach with that can provide detailed insight into cross-reactive motifs.

CHAPTER V

STRUCTURAL BASIS FOR CROSS-REACTIVITY OF A SINGLE T CELL RECEPTOR WITH TWO DIFFERENT VIRAL PEPTIDES

Author Contributions

Keith Daniels infected mice with LCMV and harvested splenocytes for further analysis.

Abstract

Although often considered to be exquisitely specific, individual T cells need to be able to recognize many different antigens in order to allow a finite pool of cells to respond to an essentially limitless variety of pathogens. One example of such cross-reactivity involves lymphocytic choriomeningitis virus (LCMV) and vaccinia virus (VV). Mice previously infected by LCMV are specifically protected from infection with VacV. Cross-reactive CD8 epitopes, from LCMV GP_{34-41} (AVYNFATM) and VacV $A11R_{198-206}$ (AIVNYANL), have been identified, but how cross-reactive T cells recognize these peptides that share only 3 identical residues is unknown. Here, we characterize a cross-reactive T cell clonotype that mediates protection between these two disparate viruses. The A22 TCR, isolated from LCMV-immune mice, recognizes GP_{34-41} bound to K^b with 2-fold higher affinity than VacV $A11R_{198-206}$ bound to K^b. Crystal structures of the respective complexes show highly similar binding modes, with extensive TCR-peptide and TCR-MHC contacts involving CDR3 from both TCR chains and similar features in the exposed peptide-MHC surfaces. In both K^b-LCMV and K^b-VacV, identical contacts are present at the shared exposed residues, P4 asparagine and P6 alanine. However, A22 does not utilize the side chain differences between GP34 P7 threonine and A11R P7 asparagine to differentiate between the two peptides. Instead, a hydrogen bond is lost in the A11R complex at P7 because the P7 asparagine is constrained to hydrogen bond with proximal K^b residues. Additionally, because K^b P3 is an auxiliary pocket formed by arginine 155, side chain differences at this position do not affect A22 binding. The propensity of K^b to shield P3, in addition to the lack of specific peptide contacts at P6 and

P7 could provide for determinants in cross-reactivity in K^b alleles. Thus, patterns of cross-recognition may vary between different MHC-I alleles due to specific MHC-peptide interactions.

Introduction

A key feature of the adaptive immune system is immunological memory. Memory T cells are reactivated upon secondary encounter of its cognate epitope. However, in some cases, unrelated peptides and pathogens can also activate memory T cells, leading to either protective cross-reactive responses or harmful pathological responses (Chen et al., 2003; Mathurin et al., 2009; Selin et al., 1998; Walzl et al., 2000; Welsh et al., 2010). T cells recognize peptide antigens T cells recognize peptide antigens bound to major histocompatibility complex (MHC) proteins using a cell-surface receptor complex (TCR) that varies clonotypically between T cells, with much of the variation present in complementarity-determining CDR1, CDR2, and especially CDR3 loops. Because of the much lower number of T cells than potential antigenic sequences, and the need for T cells to engage self-antigens during thymic selection, individual TCRs must be able to recognize and response to multiple peptide sequences. An accumulating body of work has addressed the degrees of cross-reactivity in these interactions. TCRs can tolerate many changes to a peptide sequence, and an autoimmune TCR was estimated to bind on the order of 10^6 peptides (Wooldridge et al., 2012). In studies with positional scanning peptide libraries and baculovirus display libraries, T cells were sometimes activated to a higher level by synthetic peptides than their cognate peptide (Crawford et al., 2004;

Gundlach et al., 1996; Wilson et al., 2012). Despite the cross-reactive nature of TCRs, a study which utilized yeast-displayed MHC-peptide libraries coupled multiple rounds of affinity selection and deep sequencing revealed that the optimal peptides recognized, by five mouse and human TCRs, did not deviate significantly from their known antigens (Birnbaum et al., 2014). Thus, T cells exiting selection have tonic affinity for self-MHC and only the potential to bind foreign peptides – it is not for certain that every naïve T cell will eventually bind to a foreign antigen. While TCRs may prefer to bind to an optimal peptide sequence that is found synthetically, circulating TCRs are afforded wiggle room to bind other peptide sequences.

Crystallography is often used to gain insight on TCR cross-reactive binding mechanisms. However, few examples exemplify TCR cross-reactivity, where the same TCR is ligated to different peptide-MHC complexes. Furthermore, TCRs that have been crystallized with different peptide-MHC complexes utilize different mechanisms to achieve cross-reactivity. For example, LC13 TCR was crystallized with an allopeptide, mimetic peptide (mimotope), and a viral peptide (virotope) (Kjer-Nielsen et al., 2003; Macdonald et al., 2009). Despite differences in peptide sequences and MHC alleles, LC13 CDR loops in all three complexes were oriented identically. Conversely, 2C TCR differentially docks on L^d-QL9 and K^b-dEV8 with a rotation of ~20°, resulting in differential contacts between the two complexes (Colf et al., 2007). While LC13 and 2C TCR complexes represent alloreactive responses, A6 TCR was crystallized with two foreign peptides on the same HLA-A2 allele (Borbulevych et al., 2009; Ding et al.,

1999). A6 binds to the HTLV-1 (Human T-lymphotropic virus 1) peptide, Tax, but also cross-reacts with a yeast peptide, Telp. Comparison between A6-Tax, A6-Telp complexes showed that A6 TCR docks on the MHC in a very similar orientation (Baker et al., 2012). The differences between the complexes lie in the CDR3β loop where it adopts different conformations and "tunes" to different peptide ligands.

In the case of heterologous immunity, memory responses are activated from unrelated pathogens. For example, memory T cells from mice previously infected with lymphocytic choriomeningitis virus (LCMV) protect against subsequent challenge with Vaccinia virus (VacV), producing IFNγ that helps to reduce viral titers in fat pads and liver (Selin et al., 1994, 1998). VacV challenge of LCMV-immune animals induces specific CD8+ T cell responses to LCMV epitopes (Kim et al., 2005). By examining the VacV genome for sequence similarities to known LCMV epitopes, several pairs of cross-reactive epitopes were identified (Cornberg et al., 2007). One pair of such epitopes is between GP34 (AVYNFATM) and A11R (AIVNYANL) found on LCMV and VacV, respectively (Cornberg et al., 2010).

T cell lines expanded from LCMV-immune splenocytes with the A11R peptide were also reactive to GP34, and sometimes other LCMV epitopes such as GP118 and NP205. In adoptive transfer experiments, A11R pulsed cell lines were injected into naïve congenic mice and then challenged with VacV. Mice that were adoptively transferred with a T cell line that is predominantly TCRβ 15-01 clonotype showed a log reduction in VacV titers

compared to PBS control. Furthermore, A11R specific cells are expanded in vivo by adoptive transfer of A11R-pulsed dendritic cells into LCMV-immune animal, resulting in slightly lower viral titers following VacV challenge (Che et al., 2015). Taken together, LCMV infection generates a robust CD8+ T cell response, resulting in memory T cells that are also activated and expanded by VacV epitopes. T cells that are expanded from recognition to A11R contribute to protective immunity from subsequent VacV challenge.

We previously described the molecular interactions involved in GP34-A11R cross-reactivity by generating cell lines from LCMV-immune splenocytes, followed by functional analysis of the T cell lines to peptides variants (Shen et al., 2013). When mice were infected with LCMV and assayed for IFNγ by intracellular cytokine staining (ICS) eight days post infection, splenocytes lost reactivity to both GP34 and A11R epitopes with alanine substitutions at P4. Reactivity to A11R with an alanine substitution at P7 (P7A) was not very different than wild type A11R peptide. However, reactivity to GP34 P7A was significantly diminished in comparison to wild type GP34 peptide. These data suggest that LCMV infection results in receptors with various strategies to recognize the GP34 epitope. Receptors that are cross-reactive between GP34 and A11R impinge on the P4 asparagine in both epitopes, but have varying utility at P7.

The underlying mechanisms of GP34/A11R TCR cross-reactivity are still unclear. In particular, how cross-reactivity is mediated by sharing only three amino acids (Figure V-1A) and how cross-reactive T cells differentiate between P7 residues.

While previous models of TCR cross-reactivity between foreign peptides can be explained by CDR loop conformational changes (Borbulevych et al., 2009; Mazza et al., 2007a). We previously determined the crystal structures of K^b-GP34 and K^b-A11R and suggested a molecular mimicry mechanism. Here, we isolated a cross-reactive TCR for biophysical characterization by surface plasmon resonance and x-ray crystallography to elucidate a mechanism for GP34/A11R cross-reactivity, and we show that the same TCR binds to both pMHC complexes identically, without CDR loop conformational changes between complexes.

Results

Identification of the cross-reactive TCR that recognizes both LCMV GP34 and VacV A11R

To identify a TCR cross-reactive with both GP34 and A11R epitopes, we isolated splenocytes from a mouse that had recovered from LCMV infection, and expanded the A11R-specific subpopulation by stimulation with Kb-expressing antigen presenting cells pulsed with the A11R-peptide. LCMV infection is known from previous studies to elicit populations of CD8+ T cells responsive to many LCMV epitopes (Cornberg et al., 2007, 2010; Kim et al., 2005), but after several rounds of in vitro stimulation, we observed reactivity only to GP34 and A11R peptides, and not to other LCMV epitopes such as GP118 and NP205 known from previous work to be present in the polyclonal LCMV memory population (Figure V-1B).

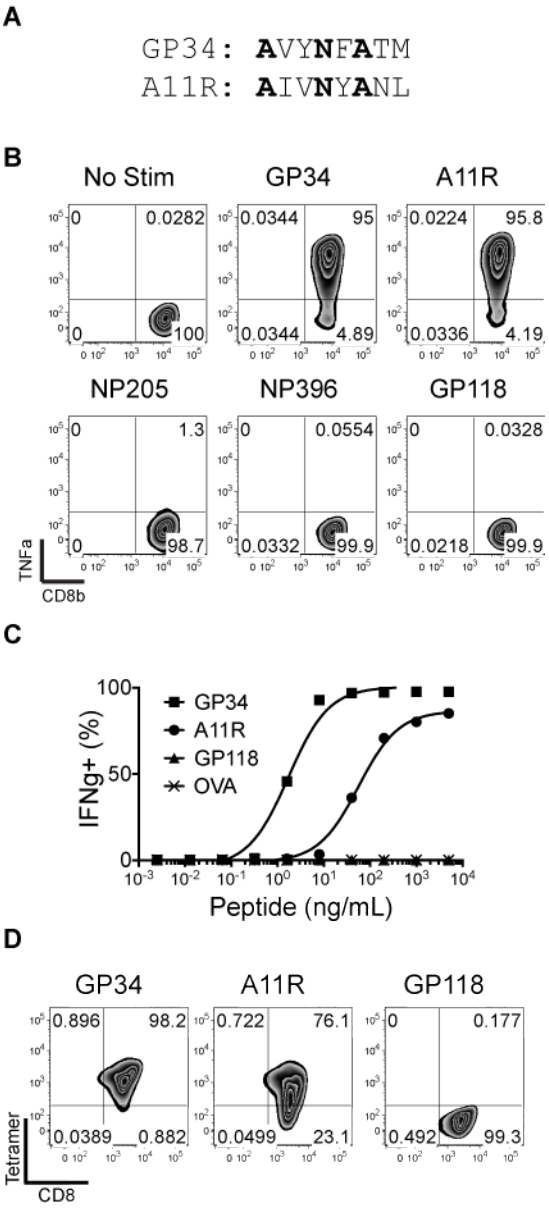

Figure V-1 Establishing a GP34/A11R Cross-reactive Cell Line from a LCMV-Immune Mouse

(A) Sequence alignment of GP34 and A11R epitopes. Identical residues are highlighted in bold. (B) A cell line was generated by co-culturing A11R pulse RMA-S cells splenocytes from an LCMV-immune mouse for multiple passages. ICCS analysis as measured by TNFα stimulated with a panel of LCMV epitopes at 1μg/mL. (C) ICCS dose response curve of GP34, A11R, and negative controls (OVA and GP118) with 5-fold dilutions from 5μg/mL to 2.5 pg/mL. (D) Tetramer staining of cell line with 0.5μM of tetramers GP34, A11R, or GP118 (control).

We evaluated this T cell line for functional responses to the GP34 and A11R epitopes, using an intracellular staining assay that measures secretion of the cytokine IFNγ (Figure V-1C). The cell line was responsive to both peptides, but more sensitive to the GP34 peptide (EC_{50} = 1.73 ng/mL) than it was to the A11R peptide (EC_{50} = 56.17 ng/mL) (Figure V-1C). No reactivity to control peptides GP118 or OVA was observed at any of the peptide concentrations tested. We confirmed the difference in sensitivity using fluorescent K^b-tetramers, which stained the T cell line more brightly when carrying the GP34 peptide (MFI = 1244) than A11R (MFI = 442) (Figure V-1D). Taken together, this T cell clone binds to GP34 tighter than A11R, which may explain the increased sensitivity in T cell activation. When probed for Vβ and Vα genes by PCR and Sanger sequencing, only one TCRα and one TCRβ sequence was detected. We will refer to this as the A22 TCR.

To verify that the observed cross-reactivity was due to the A22 TCRα and TCRβ sequences, and to rule out potential contributions in the cellular assay from differential co-receptor binding or expression of a second TCRα chain, we prepared a soluble recombinant version of the A22 TCR, and tested its binding to various K^b peptide complexes. Purified TCRs were used in a surface plasmon resonance assay (SPR) assay with biotinylated MHC-peptide complexes immobilized to streptavidin-coupled flow cells. Increasing concentrations of TCR were injected over flow cells containing K^b-GP34, K^b-A11R, a control K^b-complex (TB10 or OVA), and streptavidin-alone (Figure V-2). The maximum binding RU of each concentration was plotted against TCR

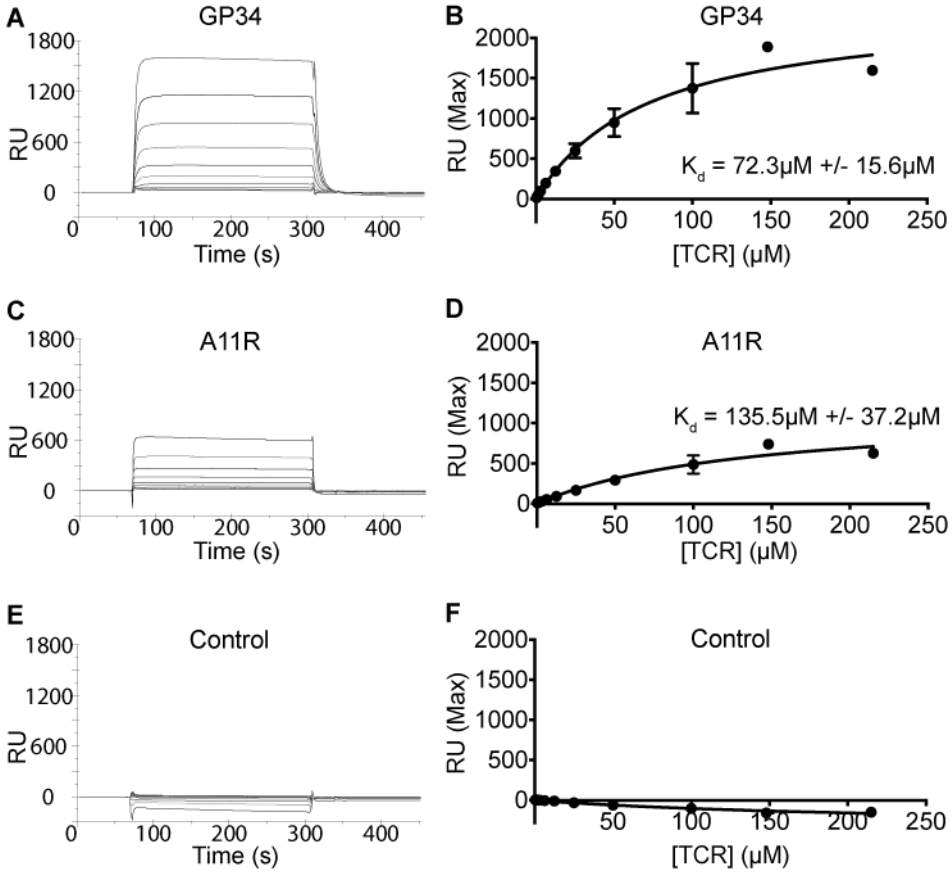

Figure V-2 Equilibrium Binding Analysis of A22 TCR with Kb-GP34 and Kb-A11R.

(A, C, E) Representative traces of Surface Plasmon Resonance analysis with background subtraction. Increasing concentrations of soluble A22 TCR was flowed over streptavidin-alone, Kb-GP34, Kb-A11R, and a Kb with an unrelated peptide (OVA or TB10). (B, D, F) Maximum binding from two independent SPR experiments are plotted against TCR concentration.

concentration and fitted to a simple two-component binding model to estimate apparent equilibrium dissociation constants K_d (Figure V-2B, D, F). A22 was observed to bind to immobilized K^b-GP34 approximately two-fold more tightly than to K^b-A11R (72.3uM +/- 15.6uM and 135.5uM +/- 37.2uM, respectively), consistent with the increased sensitivity of the T cell line to GP34 as measured by the peptide dose response assay, as well as the increased staining of GP34 tetramer relative to A11R at the same concentration.

Structures of A22 TCR, A22-K^b-GP34, and A22-K^b-A11R.

To understand the molecular basis for cross-reactive recognition of these disparate viral epitopes, we determined the crystal structures of A22 TCR alone, A22 bound to K^b-GP34 or K^b-A11R. The structure of A22 TCR alone and not bound to K^b was determined since higher resolution crystals were obtained to this than for TCR-MHC complexes (Figure V-3A, Table V-1). Despite high-resolution diffraction data, electron density was not observed for CDR1, CDR2, or CDR3 loops, except for CDR2α and the framework C'' and parts of the D strand displaced from the beta sandwich (Figure V-4). These data suggest that A22 CDR loops may have a high degree of flexibility in the absence of MHC binding.

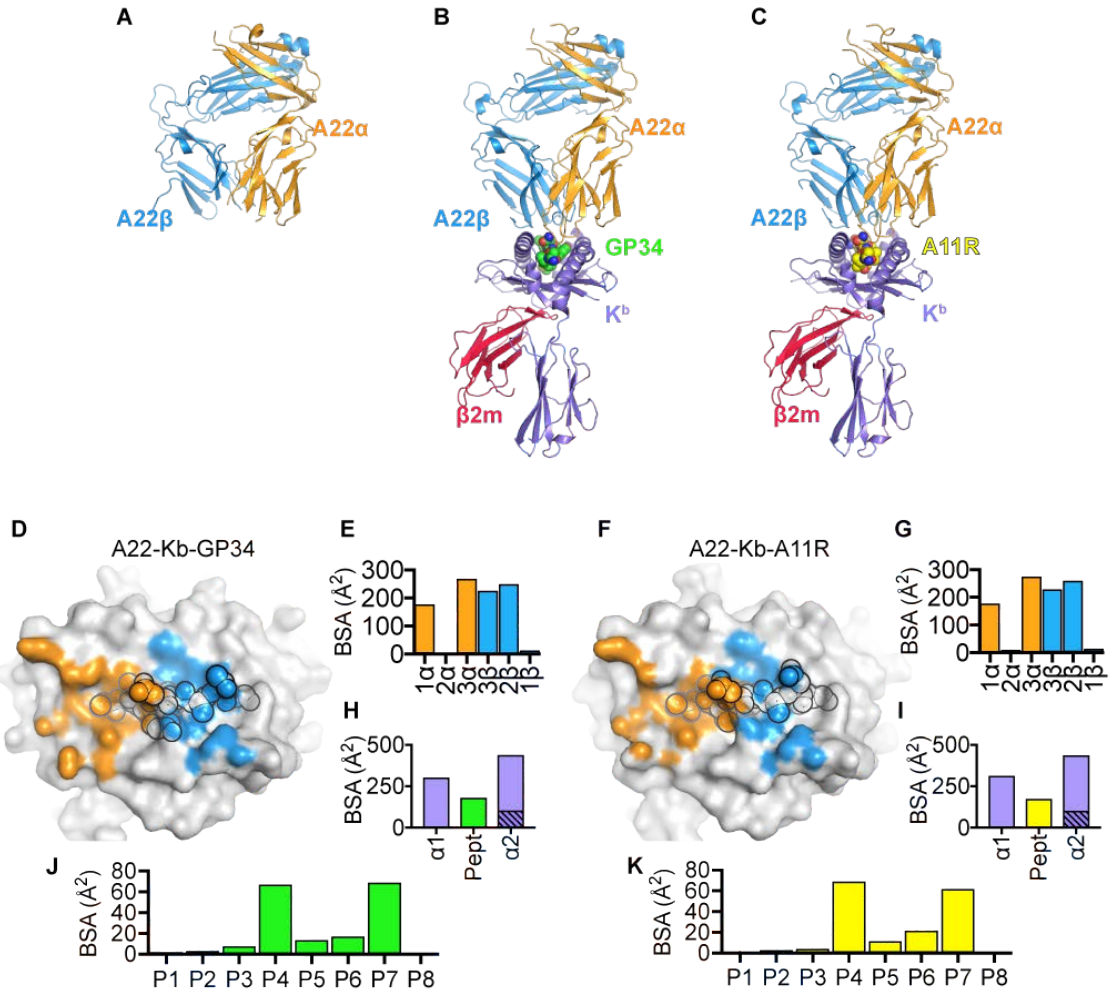

Figure V-3 Similar Binding Modes Between A22 and Kb Complexes
(A) A22 alone (blue and orange), (B) A22 in complex with Kb-GP34 (green), and (C) A22 in complex with A11R (yellow). Kb is colored purple and β2m pink. (D, F) Surface representation of contacts from A22 to pMHC complexes (Kb in grey, peptide highlighted black). Atoms within 4.0 Angstroms of A22 TCR are colored with the corresponding TCR chain colors (orange A22α, blue A22β). Buried surface area ($Å^2$) of CDR loops (E, G), pMHC complex (H,I), or peptide (J, K) in complex with Kb-GP34 or Kb-A11R respectively.

	A22	A22-GP34	A22-A11R
Source	APS LRL-CAT (31-ID)	APS LRL-CAT (31-ID)	NSLS-II FMX (17-ID-2)
Wavelength	0.979	0.979	0.979
Resolution range	26.7 - 2.152 (2.229 - 2.152)	49.62 - 3.3 (3.418 - 3.3)	49.35 - 3.356 (3.476 - 3.356)
Space group	P 1	P 61	P 61
Unit cell	41.627, 51.771, 58.012 93.078, 107.494, 107.943	249.739, 249.739, 228.727, 90, 90, 120	248.383, 248.383, 227.917, 90, 90, 120
Total reflections	89585 (8217)	1326032 (125103)	2430025 (242518)
Unique reflections	22795 (2080)	113977 (10767)	113577 (11260)
Multiplicity	3.9 (3.9)	11.6 (11.6)	21.4 (21.5)
Completeness (%)	97.01 (89.31)	93.86 (89.04)	99.72 (99.00)
Mean I/sigma(I)	22.59 (10.28)	10.67 (1.18)	11.12 (1.17)
Wilson B-factor	27.62	88.4	100.29
R-merge	0.06334 (0.1697)	0.2773 (2.425)	0.4241 (2.067)
R-meas	0.07338 (0.1963)	0.2901 (2.537)	0.4344 (2.116)
R-pim	0.03701 (0.09874)	0.08502 (0.7427)	0.09406 (0.4533)
CC1/2	0.997 (0.974)	0.994 (0.6)	0.981 (0.601)
CC*	0.999 (0.993)	0.999 (0.866)	0.995 (0.866)
Reflections used in refinement	22794 (2081)	113977 (10759)	113574 (11216)
Reflections used for R-free	1087 (90)	5656 (523)	5599 (585)
R-work	0.2306 (0.2474)	0.2616 (0.3992)	0.2591 (0.3973)
R-free	0.2462 (0.2732)	0.2753 (0.4439)	0.2752 (0.4346)
CC(work)	0.928 (0.885)	0.918 (0.665)	0.918 (0.703)
CC(free)	0.899 (0.884)	0.893 (0.581)	0.903 (0.660)
Number of non-hydrogen atoms	2999	27870	27623
macromolecules	2999	27870	27623
Protein residues	385	3571	3589
RMS(bonds)	0.004	0.005	0.004
RMS(angles)	0.98	0.95	0.95
Ramachandran favored (%)	96.99	93.13	92.27
Ramachandran allowed (%)	2.74	5.69	6.47
Ramachandran outliers (%)	0.27	1.18	1.26
Rotamer outliers (%)	1.55	2.81	3.84
Clashscore	2.25	5.36	5.84
Average B-factor	33.14	113.11	119.46
macromolecules	33.14	113.11	119.46

Table V-1 Summary of Data Collection and Model Statistics: A22, A22-Kb-GP34, A22-Kb-A11R

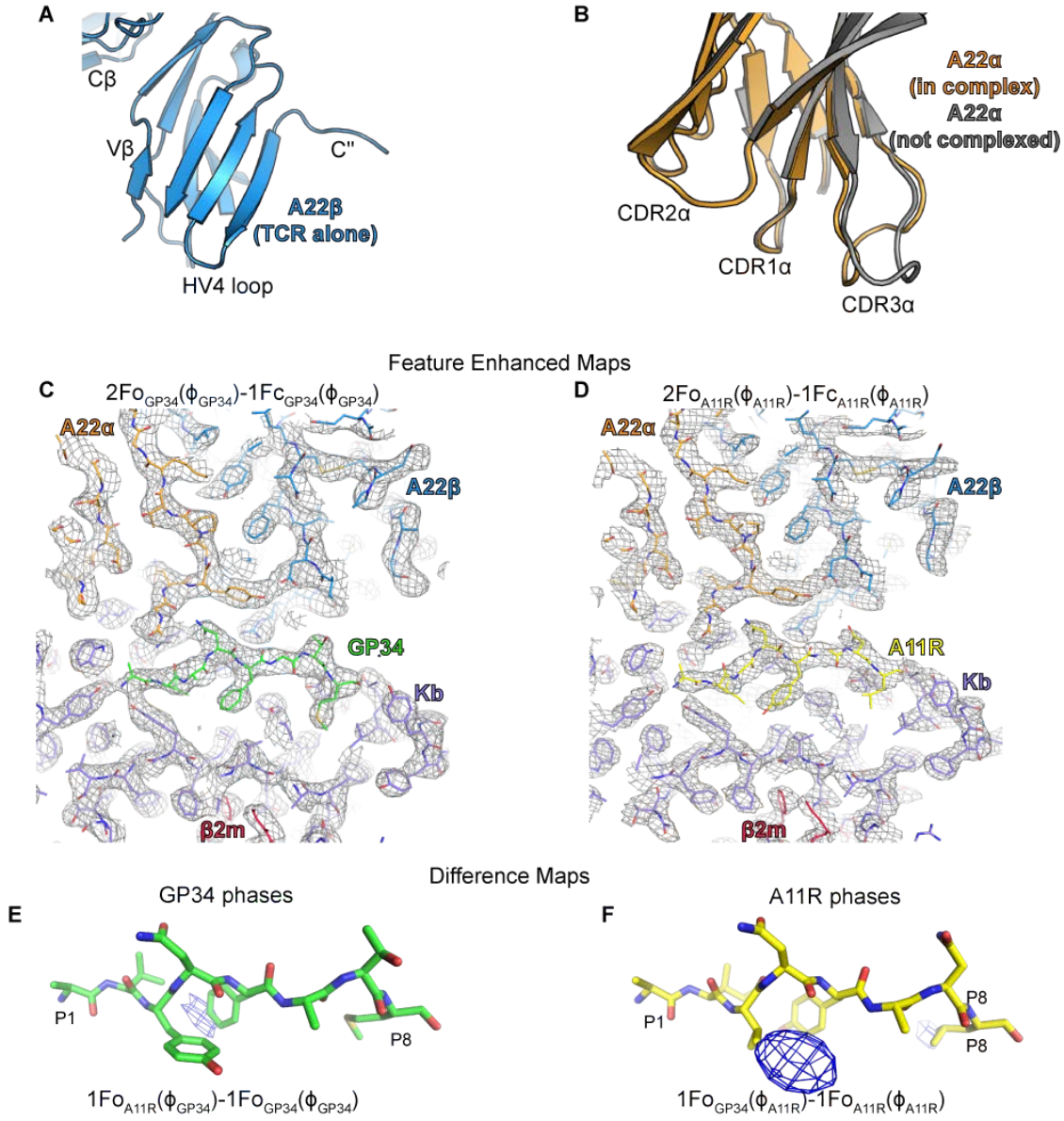

Figure V-4 Comparative Analysis of A22 Data and Models
(A) Crystal structure of A22 alone shows a hyperextended C'' strand. (B) A22-MHC crystal complexes had a 5th unligated TCR that showed an altered CDR3α loop. Model from A22-GP34 represented here. (C, D) Representative feature enchanced maps that were used for model building and evaluation. (E, F) Difference maps show that respective peptides are correctly in crystal complex datasets.

Crystal structures for the complexes of A22-K^b-GP34 and A22-K^b-A11R (Figure V-3B, C, Table V-1) were determined by molecular replacement using the A22 TCR model in conjunction with previously determined structures of K^b-GP34 and K^b-A11R (Shen et al., 2013). Both A22-K^b-GP34 and A22-K^b-A11R complexes have 4 MHC-TCR pairs and a 5^{th} uncomplexed TCR in the asymmetric unit (Figure V-4B). CDR loops were built using feature-enhanced maps (Afonine et al., 2015) (Figure V-4C, D). In each case the four copies of each MHC-TCR pair in the asymmetric unit were highly similar (RMSD = 0.28 – 1.21 Å) and the free TCR was also similar to that observed in the A11R TCR only structure (RMSD = 0.63 – 1.43). Comparison of the 5^{th} uncomplexed TCR with A22 in complex with MHC shows a slightly altered CDR3α loop conformation (~6 Å shift at glycine 97) (Figure V-4B). In the context of loop flexibility, CDR3α loop changes conformation before and after ligation. In addition to the lack of resolved CDR loops in the A22-alone dataset, these data support the notion that A22 CDR loops may be uniquely flexible.

Crystallization conditions, morphology, crystal packing, and asymmetric units were very similar between the two MHC-TCR complexes. To confirm that the crystal complexes indeed contained different peptides, Fo-Fo difference maps were calculated with both datasets and respective model phases (Figure V-4-E, F). Peak differences in the Fo-Fo maps were present at prominent size differences. A11R P3 is a valine, which is small in comparison to P3 tyrosine in GP34. Using A11R model phases, strong density is present at P3, suggesting a side chain larger than valine fits here. Surprisingly, small side chain

differences were also observed. For example, GP34 P8 methionine is a longer side chain than leucine. This difference was also observed in the difference map calculated with A11R model phases. When GP34 phases were used, density corresponding to A11R P5 tyrosine hydroxyl was observed at the tip of GP34 P5 phenylalanine. Similarly, 2Fo-Fc maps were calculated by swapping model phases to purposefully introduce phase bias (data not shown). When using GP34 model phases with the A11R dataset, density at P3 tyrosine does not extend to the side chain, despite having phase information present here. Taken together, our MHC-TCR crystals that were similar in morphology and packing contain their respective peptides.

The overall docking of A22 on K^b-GP34 and K^b-A11R are very similar between the complexes (RMSD = 0.364) (Figure V-3). We do not observe gross rotations or translations in the TCR relative to MHC in order to accommodate peptide sequence differences. In addition, CDR loops are positioned identically in both complexes. Even though we considered A22 CDR loops to be flexible, we do not observe CDR loop conformational changes between the two MHC complexes. Surprisingly, the amounts of buried surface area (BSA) involved in the interactions of MHC-TCR complex formation are relatively identical for each CDR loop, as well as individual peptide side chains.

Similar binding modes for A22-Kb-GP34, and A22-Kb-A11R

To elucidate a specific mechanism of GP34/A11R cross-reactivity, we carried out an analysis of interactions along each peptide. As usual for 8-mer peptides in K^b alleles, the

side chains of residues at the fifth and eight positions in the peptides (P5 and P8) are buried in deep pockets within the overall peptide binding groove, and the P2 and P3 residues are partially buried in shallow pockets. This leaves the P1, P4, P6, and P7 side chains available for interaction with the A22 TCR, as well the exposed surface of the peptide main chain all along the length of the peptide.

Lack of A22 Interactions at P1-P3

The A22 TCR unexpectedly focused on C-terminal side of the peptide binding site, with >90% of the peptide surface area buried by interaction with TCR involving residues at the fourth to seventh positions (P4-P7), and little to none involving residues at the first, second, or third positions (P1-P3) (Figure V3J, K). This was unexpected because previous structural characterization of isolated K^b-GP34 and K^b-A11R complexes had shown that their exposed peptide surfaces were much more similar at the amino terminal rather than carboxy terminal sides of the peptide binding groove, suggesting that cross-reactive TCR might interact near the peptide amino terminus (Shen et al., 2013). However, in the crystal structures of both K^b-GP34 and K^b-A11R complexes, A22 contacts at the amino-terminal side of the peptide binding site are minimal, and involve only CDR1a and CDR3a residues contacting MHC α1 and α2 helix residues, with little to no peptide involvement in the P1-P3 region

A22 Interactions at P4

P4 asparagine is a major contact for A22 TCR interacting with either Kb-GP34 or Kb-A11R (Figure V-3J,K), and this residue accounts for more than one third of the overall peptide buried surface area in either complex. A hydrogen-bonding network at P4 asparagine is found in both complexes (Figure V-5A, D), with the side chain amine of P4 asparagine forming a hydrogen bond network with the carbonyls of CDR3α proline 95 and glycine 97 and the main chain amide of P4 asparagine also forming a hydrogen bond with the main chain carbonyl of CDR3α proline 95. This extensive interaction and network of hydrogen bonds between CDR3α and P4N in both complexes explains why GP34/A11R cross-reactivity is particularly sensitive to P4 substitutions in either peptide (Shen et al., 2013).

A22 Interactions at P5-6

P6 alanine is found in both GP34 and A11R peptides as well. but is not a major contact site for the cross-reactive A22 TCR. P6 alanine does not form many specific interactions with A22, but does contact CDR3β arginine 96 with van der Waals interaction (Figure V-5B, E). The close contact of CDR3β to the P6 side chain explains the sensitivity of cross-reactive recognition to P6 substitution by large but not small residues (Shen et al., 2013). As noted above the P5 side chain of both GP34 and A11R peptides is buried in a deep pocket within the overall binding groove, and not available for TCR interaction, but the P5 backbone carbonyls in both complexes form hydrogen bonds with A22 CDR2β arginine 49 (Figure V-5B, E).

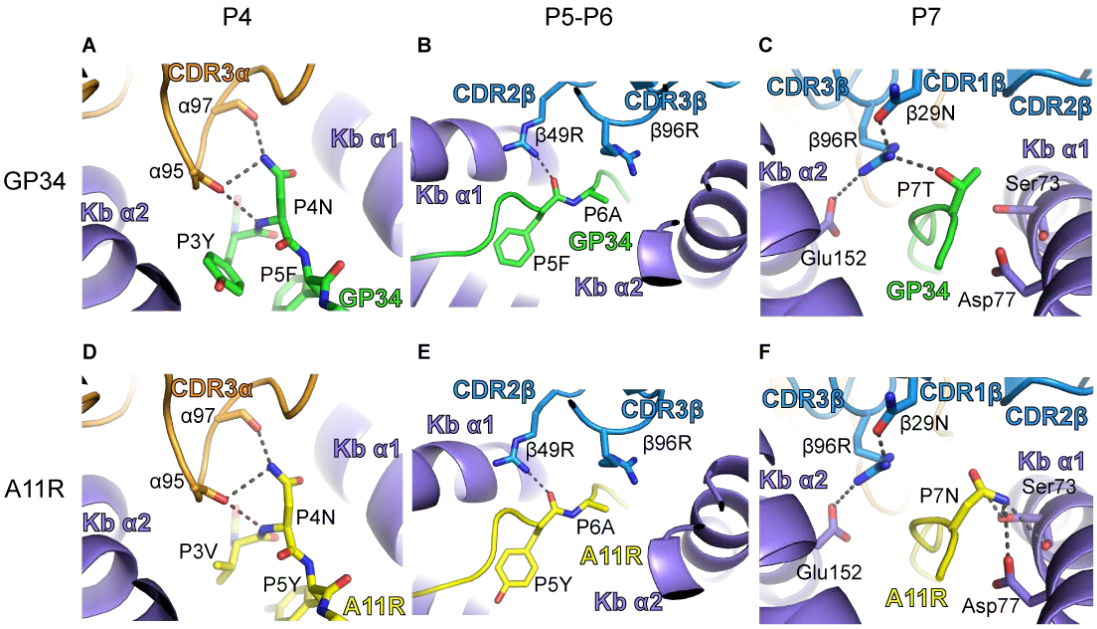

Figure V-5 Peptide Interactions are Conserved from P2-P6, but not P7.
(A, D) Conserved hydrogen bond network between A22α and P4N in GP34 and A11R complexes. (B, E) P5 backbone interaction with A22 CDR2β is conserved between the two complexes. P6 alanine is an identical residue in both peptides and interacts with Kb α2 and A22 CDR3β by van der Waals forces. (C, F) Hydrogen bond between A22β96 arginine is present in interactions with GP34 P7T, but not A11R P7N. A11R P7N forms a hydrogen bond network between Kb-73 serine and Kb-77 aspartic acid.

A22 Interactions at P7

The most apparent differences in the exposed surfaces between K^b-GP34 and K^b-A11R are present at the C-terminus of the peptide at P7 (Shen et al., 2013). Furthermore, alanine substitution of the A11R P7 residue did not abrogate cross-reactive recognition, suggesting that important MHC-peptide interaction not involve carboxy-terminal contacts. In both complexes, CDR3β arginine 96 is in the same location for both complexes and forms hydrogen bonds with CDR3β asparagine 29 and K^b glutamate 152 (Figure V-5 C, F). In complex with GP34, CDR3β arginine 96 hydrogen bonds with P7 threonine, an interaction facilitated by reorientation of the P7 Thr side chain relative to its orientation in the free Kb -GP34 complex. This interaction is missing in the A22-K^b-A11R complex, because the P7 asparagine of A11R does not reorient upon complex formation, remaining hydrogen bonded to Kb Ser 73 and Asp 77 (Figure V-5C, F). This is explained by extensive hydrogen bonding network where P7 asparagine of A11R hydrogen bonds with K^b aspartic acid 77, as well as serine 73 side chain hydroxyl and backbone carbonyl. Based on the model's Cα and Cβ positions, other asparagine rotamers will fit at A11R P7 and hydrogen bond with A22. But instead, A11R P7 interactions with K^b pull the side chain away from CDR3β arginine 96. This interaction is also present in the crystal structure of K^b-A11R without A22 TCR.

Transient Interactions at P7

The SIINFEKL (OVA) peptide variant, SIINFATL (OVA-AT), activated a previously described cross-reactive T cells line (TCL #7) (Shen et al., 2013), which was not

activated by SIINFAAL (OVA-AA). The changes between OVA and OVA-AT are due to the exposed residues in GP34 at the 6 and 7 positions. When probed with these two peptides, A22 cell line is also activated by OVA-AT but not OVA-AA (Figure V-6), indicating a possible key interaction at P7, which contrast with other cell lines that were insensitive to P7 substitutions (TCL #1 and #3) and are likely different cross-reactive clonotypes.

A22 Interactions with Kb Arg155

The amino acid residue that contributes the most buried surface area to the A22 interaction in both Kb-GP34 and Kb-A11R complexes is Kb Arg 155 (Figure V-4E, H). Arg 155 projects across the peptide binding site from the MHC α2 helix to the peptide P4 main chain carbonyl (Figure V-5A, B), in the process capping off the top of the P3 pocket and shielding the substantial peptide differences at this position (Tyr in GP34 versus Val in A11R) from TCR recognition (Figure V-7 A-F). In this orientation, Arg 155 contacts both chains of the A22 TCR, at CDR2α and CDR3β.

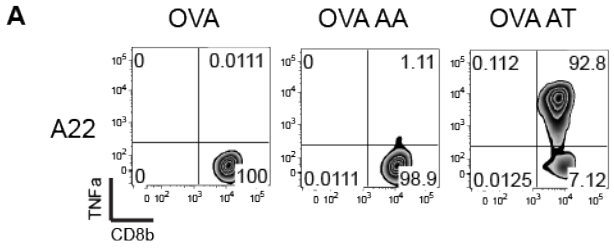

Figure V-6 Intracellular Cytokine Staining of A22 Cell Line with OVA variants. A22 cell lines were stimulated with OVA (SIINFEKL), OVA-AA (SIINFAAL), or OVA-AT (SIINFATL).

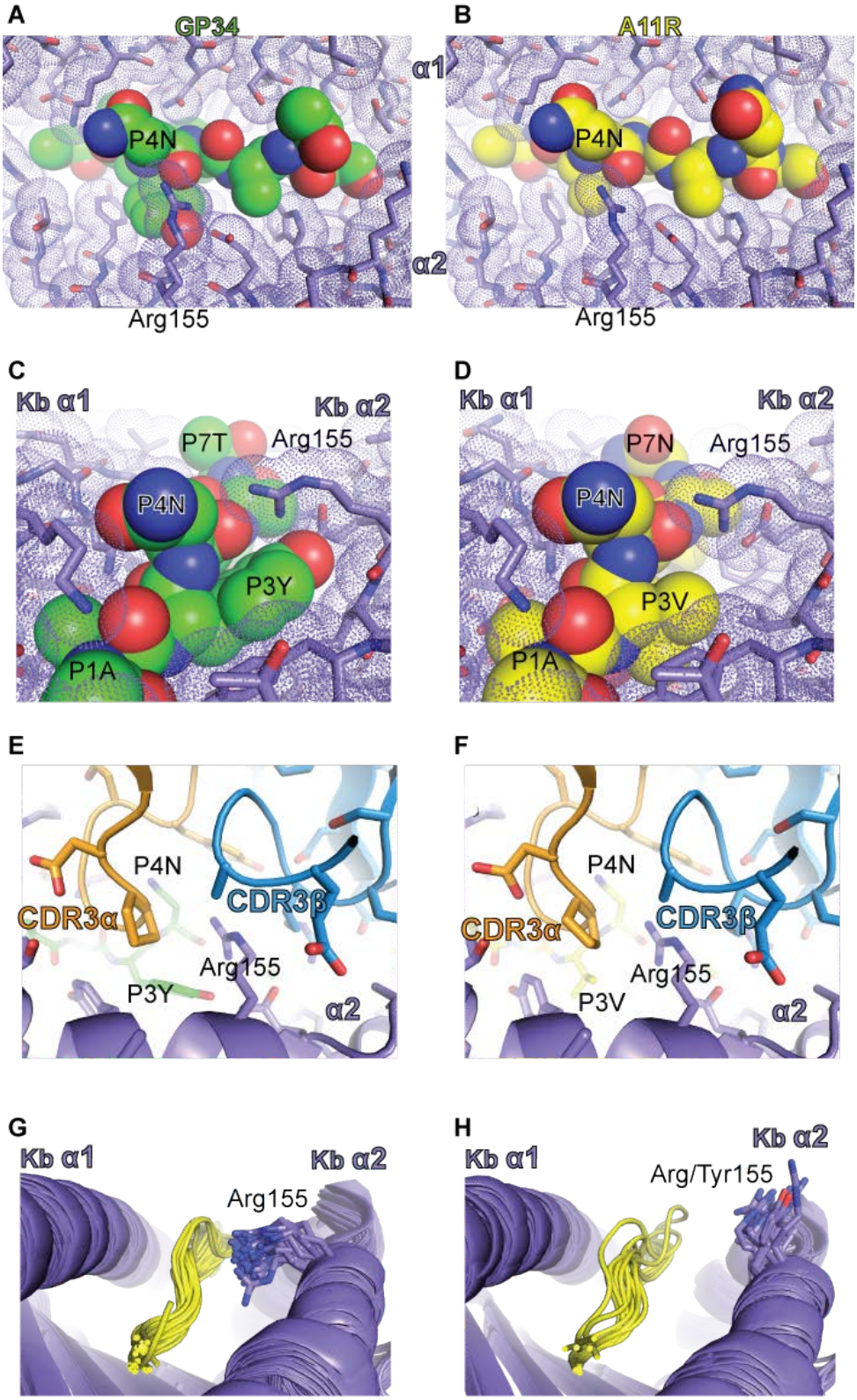
A
GP34
α1
P4N
α2
Arg155
B
A11R
P4N
Arg155
C
Kb α1
Kb α2
P7T
Arg155
P4N
P3Y
P1A
D
Kb α1
Kb α2
P7N
Arg155
P4N
P3V
P1A
E
P4N
CDR3α
CDR3β
Arg155
P3Y
α2
F
P4N
CDR3α
CDR3β
Arg155
P3V
α2
G
Kb α1
Kb α2
Arg155
H
Kb α1
Kb α2
Arg/Tyr155

Figure V-7 Kb-155 Arginine Shields Peptide Differences at P3 between GP34 and A11R.
(A, B) View of Kb-GP34 and Kb-A11R from above the peptide. (C, D) Side view of Kb-GP34 and Kb-A11R reveals a pocket over P3V in A11R that is shielded by arginine 155. (E, F) Arginine 155 is at the interface between CDR3α and CDR3β. (G, H) Alignment of 72 Kb structures by pMHC domain. 58 structures are represented in E with Kb-155 contacting the peptide and 14 in F that point away from the peptide. Two structures in F are Kb variants that have a tyrosine at 155.

Discussion

We were interested in elucidating a molecular mechanism for GP34/A11R cross-reactivity because insight into the specific interactions may have implications for determining cross-reactivity of other epitopes. These two naturally occurring viral epitopes share only three identical residues, yet LCMV infection induces expansion of A11R specific T cells, which are protective against subsequent VacV challenge. It was unclear how a single TCR could recognize both GP34 and A11R. Prior experiments detailed the reactivity of mice acutely infected with LCMV by alanine substitutions at individual peptide positions. The data suggests that many strategies are utilized to bind GP34 because many changes to the peptide were tolerated by inducing IFN secretion. Some of these strategies overlap with A11R reactivity. In particular, P4 asparagine was critical in both epitopes, but P7 substitutions had a differential effect.

Taken together, differential hydrogen bonding networks of A22 to P7 side chains may account for the differences observed by our cellular assays and SPR binding studies. The favored interaction between A22 TCR and GP34 threonine at P7 explains why GP34 peptide activates our cell line at a lower concentration. This may also explain why GP34 tetramer binds to more cells than A11R tetramer at the same concentrations as well as a higher affinity measurement of GP34 than A11R.

We hypothesized that a GP34/A11R cross-reactive TCR would alter its docking angle or that CDR loops would have alternate conformations to accommodate peptide differences at P7. Surprisingly, our GP34/A11R cross-reactive A22 TCR binds identically to both GP34 and A11R. Differential utility of the P7 side chain leads to functional and affinity differences. Among the major peptide side chain interactions with TCR (P4, P6, and P7), P4 and P6 are identical between GP34 and A11R. We show that P4 interactions with A22 involve an extensive hydrogen-bonding network with CDR3α.

P6 alanine shared between the two epitopes does not make specific electrostatic or hydrogen bonds with A22, but CDR3β 96 may be uniquely selected for side chains for van der Waal interactions with P6 alanine. Interestingly, A11R is also cross-reactive with other LCMV epitopes. GP118 (ISHNFCNL) shares P4 and P7 but has a cysteine at P6. Cell lines that were strongly GP34/A11R cross-reactive were not very GP118 reactive. Likewise, cell lines that were strongly GP118/A11R cross-reactive were not very GP34 reactive. A11R is also cross-reactive with another VacV epitope, E7R (STLNFNNL). Cell lines that are A11R/E7R cross-reactive are not cross-reactive with GP34, GP118, or NP205. One rationale may be that P6 asparagine in E7R is larger and more polar in comparison to alanine and cysteine that are present in the other cross-reactive peptides.

In the context of cross-reactivity and heterologous immunity, epitopes bound to K^b are generally more likely to be cross-reactive to another epitope than D^b. One plausible

explanation is that K^b typically binds to 8-mer peptides while D^b binds to 9-mer peptides. With a shorter peptide, there is more likelihood to chemically overlap with another peptide sequence. In addition, shorter peptides would allow less peptide surfaces for TCR discrimination. This hypothesis is supported by side chain substitutions to OVA peptide (SIINFEKL) to match P6 and P7 of GP34. Cross-reactive T cell lines are not reactive to OVA-AA (SIINFAAL), but reactivity is restored in OVA-AT (SIINFATL). These data provide two observations. First, GP34 and A11R reactivity is mediated by merely three side chains, which may be unique to K^b 8-mer epitopes. Second, A22 interactions require interactions at P7 that is not captured by crystallography. Because OVA-AA was not reactive to GP34/A11R cross-reactive cell lines, this suggests that specificity is selected for at P7. A11R P7 asparagine lacks specific interactions because of its propensity to be hydrogen bonded with K^b side chain, but alternate conformations were not captured by crystallography.

We were interested in molecular rationales for why K^b would likely present 8-mer peptides than 9-mer peptides because insight into allelic differences may provide determinants for cross-reactivity of other MHC alleles. One reason that D^b may bind to longer peptides is the presence of a ridge formed by tryptophan 73, tryptophan 147, and tyrosine 156 that might affect the peptide backbone (Young et al., 1994). These residues are serine 73, tryptophan 147, and leucine 156 respectively in K^b. Interestingly, K^b bound peptides have a slightly restricted sequence specificity at P3 which is an auxiliary pocket that is absent in many other alleles (Rammensee et al., 1993). While alanine

substitutions at P3 significantly decrease peptide binding compared to wild-type peptides, reactivity is not abrogated (Jameson and Bevan, 1992; Saito et al., 1993).

The K^b P3 pocket is formed by arginine 155 hydrogen bonding with P4 main chain carbonyl (Figure V-7A-D). In our A22-K^b structures, arginine 155 contributes the most buried surface area for any MHC side chain and likely contributes significantly to binding energetics. In only 12 out of 70 K^b models deposited in the Protein Data Bank, arginine 155 is modeled as pointing up and solvent exposed (Figure V-7G, H). Arginine 155 may play a role in shielding side chain differences at P3 for TCR discrimination, such as between GP34 tyrosine and A11R valine. However, the propensity to interact with peptide main chain may contribute to selective peptide length and specificity. TCR contacts with position 155 of MHC are highly conserved and may be important evolutionarily (Tynan et al., 2005, 2007). Contacts are maintained even when the same TCR is rotated to bind different peptides (Borbulevych et al., 2011b). Collectively, cross-reactivity may be mediated by a lack of specific MHC-TCR interactions that discriminate peptide side chains. As presented here by P3, P6, and P7, allelic differences of MHC side chains may influence the propensity for cross-reactivity.

Although prior studies describe mouse-to-mouse variability in the GP34-A11R cross-reactive lines, TRBV15-01 (Arden nomenclature Vβ 12), was frequently detected in cell lines and ex vivo (Cornberg et al., 2010; Shen et al., 2013). A cell line (L/a11r-4) bearing 97% of a TRBV15-01 clonotype was protective against VV challenge when it was

adoptively transferred into LCMV-immune hosts prior to infection (Cornberg et al., 2010). The TRBV15-01 clonotype in L/a11r had CDR3 sequences of CASSLRGGAETLY, which bears only a single amino acid substitution from the A22 TCR. Thus, A22 TCR is a representative cross-reactive clonotype that is frequently detected in LCMV-immune mice (Chapter IV).

CHAPTER VI

GENERAL DISCUSSION

Summary

The studies presented here are directed towards understanding molecular mechanisms that are involved in governing the adaptive immune system. By using x-ray crystallography, a detailed snapshot of protein-protein interactions is captured at the atomic level. The protein models presented here directly support other biochemical data or cellular assays, which was sometimes generated by my collaborators and acknowledged accordingly. While several models existed previously for MHC-peptide exchange, pMHC-TCR interactions, and TCR cross-reactivity, these models often focused on one simplistic aspect of the interactions and did not account for effects from distal regions of the interfaces. The preceding chapters presented evidence for conformational cooperatively between MHC-peptide and TCRα-TCRβ, where information about changes from distal regions is unexpectedly translated throughout the interface.

In Chapter II, I provided evidence that determinants outside of the DR1 P1 pocket contribute to stable DR1 conformations. Although it was previously shown that weak interactions at P1 decrease MHC-II peptide lifetime and increased susceptibility to DM (Pos et al., 2012), my collaborators demonstrated that peptides with a substitution at P9 peptide position could rescue weak P1 interactions and stabilize peptide off-rates. The structure that I presented of DR1-A1L9, which has a weakly associated P1 residue and a

tightly associated P9 residue, shows that the overall structure is identical to the wild-type peptide. This data supports biochemical and biophysical evidence that P1 pocket interacting residues are not the only determinants in stable DR1 conformations and DM-mediated peptide exchange. Side chains that are bound in the P9 pocket, a pocket distal to P1, also contribute to the overall stability of DR1-peptide complexes.

In Chapter III, I interrogated TCR components that may influence MHC specificity. I demonstrated that distal regions of the TCR may influence specific interactions to MHC elsewhere. The crystal structures of MHC-specific TCRs, J809.B5 and 14.C6, were solved in complex with I-Ab-3K for comparison to the parental MHC-crossreactive YAe TCR. Even though the peptide, MHC, and TCRβ sequences are constant between J809.B5 and YAe pMHC-TCR complexes, differential TCRα chain pairing alters TCRβ interactions with pMHC. Furthermore, while comparing J809.B5 and 14.C6 complexes with the same TCRα gene but different CDR3α sequences, I showed that CDR3 sequences alone could modulate CDR1 and CDR2 interactions with pMHC. This data corroborates biophysical and cellular data from our collaborators, and shows that MHC-specificity is not necessarily a product of by specific pairwise interactions between MHC and TCR. Instead, the composite TCR structure (CDR1, CDR2, and CDR3s of both TCRα and TCRβ) influences the specificity of pMHC ligands.

Lastly, I examined TCR cross-reactivity between two viral epitopes in Chapters IV and V. First, I used a deep sequencing approach to search for motifs which might confer

cross creativity upon a single T cell receptor. None were detected when the CDR3 NGS data was pooled across all TCR beta genes. After separating the CDR3 sequences by TCRβ genes and accounting for differences in abundance of each receptor, it was evident that each TCRβ gene had different cross-reactive motifs and that each motif varied in amino acid composition and CDR3 length. This supports my previous conclusions in Chapter III that the composite TCR is involved in MHC ligand specificity because each TCRβ gene encodes for different CDR1 and CDR2 sequences. I then demonstrated by x-ray crystallography that a cross-reactive TCR, A22, docks similarly to the two peptides and retains similar interactions from P1 to P6. However, A22 TCR hydrogen bonds with GP34 P7 threonine, but not A11R P7 asparagine. Although A11R P7 asparagine has the chemical make up to hydrogen bond with A22 CDR3β arginine 96, A22 arginine 96 is constrained by interactions with K^b. These data provide new evidence that TCR cross-reactivity is influenced by allelic MHC side chain interactions with the bound peptide, as well as TCR CDR loop sequences.

A thorough understanding of MHC-peptide and pMHC-TCR interactions would allow for accurate epitope prediction algorithms that could be used for therapeutic goals. However, predicting MHC epitopes and pMHC-TCR interactions is still difficult. The data presented here describe new interactions between MHC-peptide and pMHC-TCR and suggest that the composite interface between MHC-peptide and MHC-TCR are important in peptide binding and TCR binding. This view is supported by a model of conformational melding (Baker et al., 2012; Borbulevych et al., 2011c), where MHC and

TCR can cooperatively adjust for binding to proceed. In short, protein interactions are complicated.

Significance and Implications

T cell activation is one of the potent weapons for the body to fight off a pathogen's attack. There are two major reasons why it is important to understand factors that contribute to epitope selection, pMHC stability, T cell activation, and TCR cross-reactivity. The first is to identify targets for treating pathogenic infections and autoimmunity. While pathogenic agents evolve new immune evasion mechanisms, a thorough understanding of how peptides are selected and recognized can help direct new vaccination targets and strategies. Because T cells are positively and negatively selected on self-pMHC molecules, understanding TCR reactivity and cross-reactivity mechanisms can help identify therapeutic targets to prevent alloreactivity or autoimmunity. The second reason to understand the triggers of our potent immune system is to utilize the potency of the immune system for treatment of other diseases, such as cancer.

Several therapeutics approaches already take advantage of the adaptive immune system. For example, traditional vaccination strategies involve the exposure to a low dosage of live or attenuated viruses, which will result in a memory pool of immune cells. These individuals are then protected against subsequent infections of the same or similar viruses. However, introducing such large biologics with many epitopes has the potential to introduce allergenic responses (Li et al., 2014). Small peptide based vaccines are

being explored for the basis of intrinsic specificity to very targeted antigens, instead of introducing a whole virus cocktail. Currently, there are over seventy peptides approved for clinical trials (Sharfstein, 2018), but only a few have proceeded to Phase III clinical trials, most likely due to poor immunogenicity. Based on the data presented here, careful considerations should be made to avoid inducing cross-reactivity with self-peptides. While predicting reactivity and self-reactivity may aid in designing extrinsic peptide antigens as therapeutics, Chapter V of this dissertation demonstrated an unexpected MHC allelic interaction with the peptide. Currently, immunogenicity and self-reactivity are difficult to predict by peptide sequences alone. Predicting the immunogenicity of peptides may require more insight into other contextual information such as linking allelic MHC-peptide interactions and pMHC-TCR interactions.

Another therapeutic avenue takes advantage of the potency of the immune system. In contrast to vaccination strategies where an antigen is exogenously introduced and memory cell formation depends on antigen uptake, processing, presentation, and T cell recognition, exogenous TCRs are transfected directly into autologous T cells. This approach bypasses the requirement for antigen presentation by the immune system to generate a specific immune response and has the potential to utilize TCRs that are engineered with a high affinity binding to specific ligands (Li et al., 2005). TCRs directed against tumor specific antigens, such as NY-ESO-1, MART-1, and MAGE-A3, have been used in clinical trials (Chodon et al., 2014; Corrigan-Curay et al., 2014; Morgan et al., 2014; Rapoport et al., 2015; Robbins et al., 2015). While transduction of

affinity enhanced TCRs towards NY-ESO-1 was successful, TCRs against MAGE-A3 had very detrimental effects. In a study of nine patients with varying cancer types expressing MAGE-A3, two patients lapsed into coma and died. Subsequent studies determined that the TCRs transduced were highly cross-reactive with the striated muscle-specific protein, titin (Cameron et al., 2013; Linette et al., 2013). MAGE-A3 TCR was potent *in vitro*, but it unexpectedly cross-reacted with a self-protein and resulted in a massive auto-reactive response that resulted in death.

One possible reason why the anti-MAGE-A3 TCR led to cross-reactive off-target effects is because the TCR underwent many changes during *in vitro* affinity maturation. This TCR was initially selected for binding in mice expressing transgenic HLA-A2, but high affinity binding was selected *in vitro* by site-directed mutagenesis. With these changes to the TCR bypassing the normal negative selection process, it is important to include parameters that include a comprehensive panel of self-peptides and a panel of MHC alleles to assay for cross-reactivity. While pre-clinical assays did not show evidence of off-target reactivity, the data presented in this dissertation showed that subtle changes to TCR CDR loops sequences could alter MHC specificity and cross-reactivity. It is unclear if different patients' TCR repertoires or MHC alleles may have played a role in self-reactivity. Although cross-reactivity was suggested between MAGE-A3 and titin, patients showed varying pathologies and one patient fully recovered from Parkinson-like symptoms. This case study illustrates that we do not fully understand the mechanisms of cross-reactivity or epitope prediction. Because TCR repertoires are selected from an

individual's MHC haplotype and self-peptide repertoire, exogenous TCR transfusion may require detailed understanding of the interplay between MHC haplotypes and TCR responses before introducing exogenous T cell receptors into patients.

The success of CAR (chimeric antigen receptor) T cells exemplifies potency of T cells. It is an example of how to harness T cells without fully understanding the nuisances of peptide-MHC-TCR interplay. CAR-T cells are autologous T cells that are transduced with an antibody-like molecule and also carry the signaling components for T cell signaling. Upon ligand recognition, T cells are signaled for their functional outcome such as target cell killing. This approach strikes a balance between having specificity for soluble antigens and T cell effector functions. Recently, CAR-T cells were used in a B-cell malignancy clinical trial (Turtle et al., 2016) where 27 out of 29 patients achieved remission. One potential reason why CAR-T cells are thus far successful is because the antigen receptors are antibody-like and not restricted to MHC ligands. This allows for much more diversity in the sequences of receptors on the cell surface and less likelihood of cross-reactivity. Conventional TCRs may have the same potential as CAR-T cells, but MHC-restriction limits our discovery and usage of such receptors. In the absence of MHC restriction, (such as in transgenic mice that lack MHCs), T cells are expressed that recognize non-MHC molecules (Tikhonova et al., 2012).

As we move towards personalized medicine, it is prudent to consider an individual's MHC and TCR haplotype. The data presented here suggests that MHC alleles play a role

in ligand reactivity and cross-reactivity. In the same vein, an individual's TCR alleles may also play a role in ligand recognition and is an avenue that remains to be investigated

Limitations and Future Directions

One criticism of x-ray crystallography is that the data represents a static view, which may be different than relevant dynamic interactions in solution or *in vivo*. The crystal structure of MHC bound to peptide presented in Chapter II is an excerpt from a co-authored manuscript. The larger study followed the crystal structure up with several additional assays including conformational specific antibodies, biochemical assays, and NMR spectroscopy. In particular, NMR spectroscopy revealed that the significant changes between DR-A1 and DR-A1L9 complexes are present at pocket 1, despite having the same P1 residue in both complexes. This directly supports the structure presented here, that A1L9 is a stable complex due to the L9 substitution and samples a DM-resistant conformation in comparison to DR1-A1.

MHC-TCR interactions are often difficult to measure quantitatively. Some groups use alanine scanning substitutions or fluorogenic probes to quantify energetic requirements at the interface. However, changing residues at the interface may unpredictably affect ligand specificity that was previously constrained by positive and negative selection of T cells. The argument presented here supports a model of conformational melding between peptide-MHC and MHC-TCR. NMR would be an ideal method to test this hypothesis,

but the large size of MHC-TCR complexes and relatively weak binding affinities would provide complications for a reliable and conclusive analysis.

Hence, the understanding of MHC-TCR interactions has relied heavily on MHC-TCR crystal structures. Here, we isolated and determined the structures of two self-tolerant MHC-TCR complexes. These receptors were isolated from TCRβ transgenic mice in order to study TCRs that pair with other TCRα chains. This is a unique study in the field where MHC, peptide, and TCRβ are shared. Additionally, the complex of 14.C6 TCR uses an identical TCRα chain as J809.B5, but carries a different CDR3 sequence. One limitation of this study is that the parental TCR used to generate the TCRβ transgenic mice, YAe62, was isolated as self-reactive TCR. Although only the YAe62 TCRβ chain is used as a transgene, the TCRs isolated from these mice may intrinsically be biased towards self-reactivity and have affects that are unaccounted for. The crystal structure of the YAe62-pMHC complex was slightly tilted, where the TCRβ chain made many more contacts with MHC than TCRα. One potential avenue would be to further characterize TCRs from YAe62 TCRα transgenic mice, instead of YAe62 TCRβ, to ascribe a role of the α chain in MHC self-reactivity. Another avenue of interest is to characterize TCRα-TCRβ chain pairing in other systems. Nonetheless, the overall conclusions from this study, that the composite TCR is involved in ligand recognition, are supported by the conclusions of GP34/A11R repertoire analysis.

GP34/A11R cross-reactive motifs were identified after a logical process to parse NGS data. Since CDR1 and CDR2 sequences vary between TCR genes, different genes exhibit different cross-reactive motifs. Compared to other studies, the approach in this study is unique in the method by which we separated our data. Our study is also unique in that CDR3 sequences were weighted accordingly, increasing the weight of a sequence based on how many times it was read by NGS. Many TCR NGS analyses do not use these corrections, which may be a more accurate representation of the TCR repertoire *in vivo*.

Many studies do not correct for the number of times a sequence is observed because PCR steps can amplify primer biases. Additionally, it is difficult to obtain large enough sample quantities *ex vivo* for NGS analysis. Therefore, RNA is typically used as starting material for TCR NGS. We did not use this approach because RNA copy numbers may vary from cell to cell and introduce bias towards different receptors. Instead, we used genomic DNA to circumvent this issue. The NGS platform that was used also includes controls for quantification such as internal synthetic standards and a bioinformatics correction for PCR bias. A potential pitfall in our approach is comparing datasets between samples. A rarefaction analysis, which attempts to estimate the true diversity from a sample, showed different diversity measurements between different datasets. While it is still unclear of what type of analyses best represents TCR repertoires, the rarefaction analysis suggests that we might not be capturing the full breadth of a mouse repertoire. However, this may also be due to the fact that we pooled 15 mice, with each

pool containing a different set of mice, and each mouse possessing its own private TCR repertoire, a subset of receptors that are not shared between other mice.

Our NGS dataset showed that many A11R-reactive receptors could be isolated from LCMV-immune mice. Interestingly, only a subset (~18%) of A11R-reactive cells is also reactive to GP34. Because the A11R-reactive population is only detectable after LCMV infection, by definition, A11R-reactive cells are cross-reactive with a LCMV epitope. To understand other A11R cross-reactive patterns, future experiments would include tetramer staining with LCMV epitopes, such as E7R or GP118, followed by TCR NGS as performed here. I would expect that different cross-reactive motifs would be apparent that are different from A11R/GP34 cross-reactive motifs.

As of writing, our understanding of the cross-reactive TCR repertoire is limited to TCRβ sequencing. It is difficult to capture paired αβ sequences with the current breadth of TCRβ sequencing. As technology advances, a key avenue of research will be investigating how TCRα-TCRβ pairing influences cross-reactivity.

Lastly, the cross-reactivity observed between GP34/A11R is unique to K^b. As presented in Chapter V, cross-reactivity between GP34 and A11R is mediated constrained peptide interactions between P7 asparagine and K^b allelic side chains. Compared to other MHC complexes, K^b epitopes were observed to be more susceptible to cross-reactivity. While we captured complexes that bind similarly between the two peptides, other mechanisms

could potentially mediate cross-reactivity in different receptors, such as alternate docking angles or CDR loop conformational changes. Determining the structures of other cross-reactive TCRs in other MHC alleles will be of interest. I note that unique MHC allele specific interactions with the peptide may contribute to cross-reactivity, such as K^b arginine 155 and interactions at P7, but it remains to be seen if it is a general TCR cross-reactive mechanism.

APPENDIX A

MATERIALS AND METHODS

DR1-A1L9 (PDB: 4OV5) - Crystallization Condition, Data Collection, and Structure Determination

Conditions: 4 °C in 4% PEG 4000, 10% ethylene glycol, 100mM sodium acetate (pH 5.0).

Cryocondition: Single 100 x 100 x 300 um crystals were transferred to cryo-solution containing 30% ethylene glycol and flash-cooled in liquid nitrogen.

Beamline: National Synchrotron Light Source X29 beamline.

Data Indexing and Scaling: HKL2000

Phasing: Molecular replacement with previously solved structure of DR1-A2 (PDB:1AQD) using Phaser (CCP4i). Residues within 4 Å of the peptide were removed.

Crystal Composition: P 21, 6 molecules/ASU

Model Building: An initial model with ambiguous density was trimmed using a composite omit-maps (CNS). Models were built into 2Fo-Fc maps.

Refinement Strategy: XYZ positional refinement, Individual B factors

809.B5 TCR and IAb-3K (PDB: 3RDT) - Crystallization Condition, Data Collection, and Structure Determination

Conditions: 1:1 molar 809.B5 TCR and IAb-3K mixed at 10mg/ml using hanging drop vapor diffusion method at room temperature. 12% PEG 4000, 100 mM sodium citrate, and 100 mM sodium cacodylate, at pH 5.0

Cryocondition: Well solution containing 25% (w/v) glycerol and flash-cooled in liquid nitrogen.

Beamline: National Synchrotron Light Source X29 beamline.

Data Indexing and Scaling: HKL2000

Phasing: Molecular replacement with previously solved structure of YAe-IAb-3K (PDB:3C60) using Phaser (CCP4i). Residues within 5 Å of the peptide were removed.

Crystal Composition: C 2, 1 molecule/ASU

Model Building: An initial model with ambiguous density was trimmed using a composite omit-maps (CNS). Models were built into 2Fo-Fc maps.

Refinement Strategy: XYZ positional refinement, Group B factors, TLS

14.C6 TCR and IAb-3K (PDB: 4P5T) - Crystallization Condition, Data Collection, and Structure Determination

Conditions: 1:1 molar 14.C6 TCR and IAb-3K mixed at 10mg/ml using hanging drop vapor diffusion method at room temperature. 100 mM sodium cacodylate, 200 mM sodium citrate (pH 5.5) 10% (w/v) PEG 4000 and 0.5% (w/v) n-octyl-b-D-glucoside.

Cryocondition: Well solution containing 25% (w/v) glycerol and flash-cooled in liquid nitrogen.

Beamline: National Synchrotron Light Source X29 beamline.

Data Indexing and Scaling: HKL2000

Phasing: Molecular replacement with previously solved structure of B5-IAb-3K (PDB:3RDT) using Phaser (CCP4i). Residues within 5 Å of the peptide were removed.

Crystal Composition: P 21, 2 molecule/ASU

Model Building: An initial model with ambiguous density was trimmed using a composite omit-maps (CNS) and RESOLVE map. Models were built into 2Fo-Fc maps.

Refinement Strategy: XYZ positional refinement, Group B factors

A22 TCR - Crystallization Condition, Data Collection, and Structure Determination

Conditions: 5-10mg/ml using sitting drop vapor diffusion method at room temperature. 100 mM sodium acetate pH trihydrate pH 4.0, 8% (w/v) PEG 3350

Cryocondition: Well solution containing 30% (w/v) PEG 400 and flash-cooled in liquid nitrogen.

Beamline: National Synchrotron Light Source X29 beamline.

Data Indexing and Scaling: HKL2000

Phasing: Molecular replacement with a composite TCR structure with 3C5Z and 3UTT TCR chains using Phaser (CCP4i). Initial model was trimmed using Scultor.

Crystal Composition: P 1, 1 molecule/ASU

Model Building: An initial model with ambiguous density was trimmed using a composite omit-maps (Phenix). Models were built into 2Fo-Fc and Feature enhanced maps.

Refinement Strategy: XYZ positional refinement, individual B factors, TLS

A22-Kb-GP34 - Crystallization Condition, Data Collection, and Structure Determination

Conditions: 5-10mg/ml using sitting drop vapor diffusion method at room temperature. 100 mM sodium cacodylate pH 7.0, 1.5% tacsimate pH 4.0, 4% (w/v) PEG 3350

Cryocondition: Well solution containing 30% (w/v) PEG 400 and flash-cooled in liquid nitrogen.

Beamline: Advanced Photon Source 31-ID.

Data Indexing and Scaling: aimless

Phasing: Molecular replacement with a composite TCR structure with 3TID (Kb-GP34) and partially built A22 TCR using Phaser (Phenix).

Crystal Composition: P 61, 4 MHC:TCR molecule/ASU, 1 TCR molecule/ASU

Model Building: An initial model with ambiguous density was trimmed using a composite omit-maps (Phenix). Models were built into 2Fo-Fc, Feature enhanced maps, and NCS maps.

Refinement Strategy: XYZ positional refinement, group B factors, TLS, NCS restraints

A22-Kb-A11R - Crystallization Condition, Data Collection, and Structure Determination

Conditions: 5-10mg/ml using sitting drop vapor diffusion method at room temperature. 100 mM sodium formate pH 5.0, 8% (w/v) PEG 3350

Cryocondition: Well solution containing 30% (w/v) PEG 400 and flash-cooled in liquid nitrogen.

Beamline: NSLS-II FMX 17-ID-2

Data Indexing and Scaling: aimless

Phasing: Molecular replacement with a composite TCR structure with 3TIE (Kb-A11R) and partially built A22 TCR using Phaser (Phenix).

Crystal Composition: P 61, 4 MHC:TCR molecule/ASU, 1 TCR molecule/ASU

Model Building: An initial model with ambiguous density was trimmed using a composite omit-maps (Phenix). Models were built into 2Fo-Fc, Feature enhanced maps, and NCS maps.

Refinement Strategy: XYZ positional refinement, group B factors, TLS, NCS restraints

Structural Analysis

Contacts were identified with NCONT using a 4 Å cut off.

Buried Surface Area and Hydrogen bonds were identified with PISA webserver.

Structural alignments were done in PyMol or LSQMAN.

NCS copies were moved with PyMol and Coot.

B factors were reset to Wilson B factor prior to initial refinement using Moleman2.

Next Generation Sequencing for TCRβ

Isolated splenocytes were enriched and untouched using MACS CD8+ T cell Isolation Kit prior to cell surface staining with antibodies and tetramer.

FACS sorted cells were left unfixed for DNA isolation.

Genomic DNA was isolated using Qiagen DNeasy Blood & Tissue Kit and QIAamp DNA Micro kits.

DNA quantification was accomplished with Quan-IT PicoGreen assay

NGS amplication and libraries were generated according to Adaptive Biotechnologies immunoSEQ kit.

Data analysis was carried out using Python. The Pandas library was used to reindex sequences by V gene and CDR3 sequence. Plotly library was used to plot scatter plot graphs.

PCR reaction (adapted from Sambrook Molecular Cloning 3rd Ed, Section 8.21)

- 50ul reaction
- 200uM each dNTP
- 50 units/ml of Taq
- 1uM of fwd primer
- 1uM of rev primer
- 1 ng of plasmid

Example Reaction:
Primers diluted to 20uM (2ul of 100uM +8ul of dH2O)

- 2.5ul of primers
- 0.5-1ul of Taq
- <1ul of plasmid
- 1ul of dNTP
- bring volume to 50ul with dH2O
-

Taq Master Mix (30ul per reaction)

- 13ul dNTP
- 65ul of 10x Taq buffer
- 312ul of dH2O

EtBr Gels

- Stock EtBr @ 10mg/ml
- Dilute 1:40, to get 250ug/ml (25ul of 10mg/ml EtBr + 975ul of TAE buffer)
- Agarose gel should have 0.5 ug EtBr per ml of agarose
- Add 100ul of 250ug/ml EtBr per 50ml of agarose

DNA ladders from NEB

- load 0.5ug per well, NEB stock is 500ug/ml
- for 120ul, mix 12ul 100bp + 88ul dH2O + 20ul 6x loading buffer

Intracellular Cytokine Stain and Tetramer Stain of T cell Lines
Adapted from Keith Daniels/Mike Brehm/Zu Shen

Reagents

	Company	**Cat. Number**	**Dilution**
GolgiPlug	BD	51-2301KZ	0.2ul/200ul
IL-2 Diluted (Neat = 0.004ul of 1ug/ml 10U/ml)			0.4ul/200ul
CD3			0.1ul/200ul
Lympholyte	CedarLane Labs	MCL5035	
Fc Block CD16/CD32 (unlabeled)	BD	553142 (clone 2.4G2)	1:100
CD44 (FITC)	BD	553133	1:400
CD8 (PerCP Cy5.5)	Biolegend	126610 (clone YTS156.7.7)	1:200
Live Dead (Aqua)	BD	L34957	1:300
TNFa (PE)	Biolegend	506306	1:200
IFNg (APC)	BD	554413	1:1000
FACS Buffer (PBS w/o Mg++ or Ca++; 1% FCS; 0.09% NaN_3)			

Protocol
(Specific peptide stimulation)

Stimulation

1. Plate cells at 100 ul per well in 96 well U-bottom plate. (Final volume will be 200 ul)
2. Make master mix of IL-2 and Golgi plug for each well.
3. In free wells, plate master mix to add peptide for distribution. Add 1ul of peptide for each well distributed. (i.e. 3 ul for 300 ul of mix) Distribute 100 ul to cells.
4. Incubate at 37 degrees C for 4.5 hours

Fc Block

5. Wash 1x with FACS buffer. Add 100 ul FACS buffer; centrifuge 5 min at 300 g (1200 RPM).
6. Block Fc receptors with Fc Block antibody. 5 min at 4 degrees C in 100ul FACS buffer.
7. Wash 1x with FACS buffer. Add 100 ul FACS buffer,

Surface Antigens

8. Make master mix of surface antigens in 100 ul per well.
9. Add to cells and incubate 20 minutes at 4 degrees C.
10. Wash 2x with FACS buffer.

Permeate Cells

11. Resuspend in 100/ul Cytofix/cytoperm for 20 min at 4 degrees C
12. Wash 2x in 1x *Perm/wash solution*

Intracellular Cytokines

13. Make master mix of intracellular cytokines in 100 ul per well.
14. Add to cells and incubate 25 minutes at 4 degrees C.
15. Wash 2x with 1x Perm/wash.
16. Resuspend in 150 ul of FACS buffer.

Tetramer staining Protocol

Cell Clean Up

1. Resuspend cells in 5 ml increments of media, 1-2 million cells/ml.
2. Add 5 ml of Lympholyte-M to 15 ml conical.
3. Slowly add 5 ml of cells (5-10 million cells) per 5 ml of Lympholyte-M.
    a. Add cells at an angle and very slowly, not to disrupt the interface.
4. Centrifuge 30 min at 900 g (2000 RPM), 25 degrees C.

Wash Cells After Lymphyolyte

5. Discard top 3 ml of media.
6. Pipette up the interface, media, and up to 2-3 mls of Lympholyte-M.
7. Wash 2x by adding 10 ml of media and spinning at 1200 RPM for 5 minutes.
8. Resuspend in media and count with trypan blue.
9. Plate about 0.5×10^6 cells per well.
    a. Resuspend at 5×10^6 per ml, then plate 100 ul per well.

Stain Cells with Tetramer and Surface stains

10. Add Tetramer stains for 45 minutes at 4 degree C.
11. Add surface stains for 25 minutes at 4 degree C.
12. Wash 1x with FACS buffer.
13. Add 100 ul Cytofix, neat from bottle for 5 min.
14. Wash 1x with FACS buffer.
15. Resuspend in 150 ul FACS buffer.

SDS Gel Loading Buffer
(Adapted Protocol from Guoqi)

Reagents Needed:
1M Tris HCL, pH 6.8; FW=157.6
DTT; FW = 154.25
Bromophenol blue; FW = 670
Urea; FW = 60.04

1x solution contains the following concetrations:
125mM Tris, pH 6.8
10% glycerol
2% SDS (w/v)
100mM DTT
0.002% bromophenol blue

4x solution contains the following concentrations:
500mM Tris, pH 6.8
40% glycerol
8% SDS (w/v)
400mM DTT
0.008% bromophenol blue

To make a 4x solution:

- 5ml 1M Tris, pH 6.8 (same pH as stacking gel)
- 4ml glycerol
- 0.8g SDS
- 617mg DTT
- 8mg bromophenol blue

Bring to 10ml. Store 500ul aliquots in -20 degrees C.

→To make a urea gel loading buffer, use 5ml of 1M Tris, pH 6.8 made with 8M urea.

Cast SDS Page Gels
(Adapted protocol from Guoqi)

	X%	4%	12%
30% Acrylamide/Bis	3.3(X%)=**A** ml	3.3ml	40ml
0.5M tris, pH 6.8	-	6.3ml	-
1.5M Tris, pH 8.8	25ml	-	25ml
10% SDS	1.0ml	250ul	1.0ml
dH2O	73.5-**A**	15ml	33.3ml
TEMED	50ul	25ul	50ul
10% APS	500ul	125ul	500ul
Final Volume	**100ml**	**25ml**	**100ml**

Reagents will make 10 gels in Biorad casting system.

1. Mix reagents for 12% resolving gel, then pour into gel apparatus.
    a. Add TEMED and APS last.
2. Carefully Layer top with saturated Tris-butanol solution. (1:4 stacking buffer Tris:butanol) (Water can also be used instead of butanol)
3. When resolving gel has solidified, pour off Tris-butanol. Rinse gently with dH2O.
    a. It takes about 1 hour to solidify in Biorad apparatus for 10 gels. It may be shorter with less gels and a different apparatus.
4. After ~1 hour, mix stacking gel, layer on top of resolving gel.
5. Quickly put combs in between gel plates.
6. When gels have solidified, wrap gels individually with moist paper towels and store in 4 degrees C.

Biotinylation Yield Test for MHC-I Complexes

Protocol modified from Jennifer Stone

Procedure**:**

1. Set up reducing SDS-PAGE using a 12% polyacrylamide gel.
2. Assemble the following samples in reducing SDS loading buffer and boil your samples. Make sure to add enough reducing SDS loading buffer to account for the addition of streptavidin or PBS in step 3. Samples b-d should be brought to the same final volume.
    a. Protein MW standards
    b. 2 μg Biotinylated MHC class I + PBS or other buffer
    c. 2 μg Biotinylated MHC class I + 2 μg Streptavidin
    d. 2 μg Streptavidin
3. Cool all boiled samples to at least room temperature. Then add an excess of streptavidin (2 μg) to the samples and an equivalent volume of PBS or similar buffer to the unbiotinylated monomer sample. Allow all samples to sit at least 10-15 minutes to allow for the streptavidin-biotin interaction to occur.
4. Run samples on gel (140V for 1 hour), then stain with coomassie. I usually rinse the gel after coomassie staining with deionized water and then add the destain.
    a. Note, when destaining, tie two kimwipes (single knot) and immerse in the destaining solution. The kimwipes will bind up the coomassie.
5. The band corresponding to the biotinylated heavy chain of your MHC class I should disappear in sample C, and larger bands will appear indicating the presence of streptavidin-MHC oligomers.
6. The yield of biotinylation can be determined by determining the fraction of MHC class I heavy chain remaining after streptavidin was added.

Generation and Maintenance of Antigen-Specific T cell Lines

Protocol modified from Michael Brehm and Zu Shen

Materials*:*

- RMA cells (Roche virus transformed lymphocyte derived from B6 mice), expressing K^bD^b
- T cell media
  - RPMI, 10% FCS, 1% P/S, 5 mL HEPES, 5 mL L-Glut, 5 mL nonessential amino acids, 5 mL sodium pyruvate, 2.5 mL 0.01M BME
    - All purchased from Invitrogen (except BME)
- RP10
  - RPMI, 10% FCS, 1% P/S, 5 mL HEPES, 5 mL L-Glut
- 1 uM peptide (VV-A11$R_{198-205}$, LCMV-GP_{34-41})
- Filtered tips to prevent cross contamination of peptide
- 0.8% Sterile Ammonium Chloride or Sigma RBC lysis buffer (Cat #R7757)
- 12 well plates (4 mL well volume ideal)
- BD T-Cell Culture Supplement without conA (IL-2 culture supplement), rat (BD Cat# 354115)

Procedure*:*

1. Spin down RMA cells for 5 minutes 1200 rpm and resuspend in T cell media to 10 million cells/mL
2. Pulse cells with peptide by adding 5 uL (200 ug/mL stock) peptide for every 10 million cells/mL
3. Incubate for 1 hour in CO_2 incubator, leaving 50 mL conical caps loosened
4. Irradiate pulsed cells for 3000 rads
   a. Wash irradiated cells with 25 mL T cell media and resuspend in 1 million/mL
5. Mash up spleens with 3 mL syringe top (inside of mesh) and add 5 mL of 0.8% Ammonium Chloride OR Sigma Lysis buffer for 5 minutes
   a. Add 25 mL of media to stop the lysis and spin down 1200 rpm for 5 minutes, discard supernatant
6. Count and resuspend splenocytes in 3.33 million/mL
7. Want a 10:1 ratio of splenocytes:stimulator RMA cells
8. Add 1 million cells (1 mL) of pulsed, irradiated cells to culture and 10 million (3 mL) of splenocytes to the culture
9. Check culture every 3-4 days, add 1 mL of T cell media if necessary

Restimulation procedure:

1. Spin down stimulator RMA for 5 minutes 1200 rpm and resuspend in T cell media to 10 million cells/mL
2. Pulse cells with peptide by adding 5 uL (200 ug/mL stock) peptide for every 10 million cells/mL
3. Incubate for 1 hour in CO_2 incubator, leaving 50 mL conical caps loosened
4. Irradiate pulsed cells for 3000 rads
    a. Tighten cap on conical tube and place tube in irradiator
        i. To operate irradiator, turn on, put samples in and set time and make sure the timer starts counting upwards
        ii. Set to automatic
        iii. When completed, follow procedure to exit (COMMAND, ARM, YES)
    b. Wash irradiated cells with 25 mL T cell media and resuspend in 1 million/mL
        i. Can use between 500,000 and 1 million stimulator cells per well
5. Harvest the cells by taking up the full 4 mLs in each well, and then pipet down in a circular motion to get all of the cells, do this twice for each well and then add to a 50 mL conical
6. After harvest, add 1 mL PBS and put back into incubator for 5-10 minutes while harvesting other plate(s)
    a. Take 4 mLs of the harvested cell mixture and use that to harvest the 1 mL of PBS added + cells and add to 50 mL conical
    b. Adding PBS will help remove the cells that are stuck to the bottom of the well, this is more important for latter passages
7. Spin down harvested cells for 5 minutes at 1200 rpm
8. Reseed cell lines based on density, which is determined visually (seed 1:2 or 2:3 for the first 3 passages and 1:3 or 1:4 for later passages)
9. Resuspend pelleted cells using T cell media containing 10% T stim without ConA
    a. Make cocktail with T cell media so that cells can be resuspended into 3 mLs each well
10. Add 1 mL of pulsed RMA stimulators (between 500,000-1 million) with reseeded T cells and incubate for 3-4 days

Inclusion Body Preparation for MHC-I Subunits
Protocol modified from Mia Rushe

Procedure (for pellet from 10 L culture):

1. Spin down cells at 5000 x g—collect supernatant into a container, sterilize with 1% Wescodyne for 10-20 minutes, and then dump down the sink.
2. With a rubber spatula, re-suspend fresh bacteria into a single plastic container with ~200 mL Sucrose Solution.
3. Chop the solution briefly in a homogenizer or polytron. **Do not sonicate!
4. Most of our cells are not $pLys^{s}$, so add 1 mg dry lysozyme per mL suspension (0.2 g) and stir for 10 minutes. If cells are $pLys^{s}$, just stir for 10 minutes.
5. While stirring, add 500 mL Deoxycholate-Triton Solution. Solution will become very viscous due to cell lysis and DNA release.
6. Add 1 mL of 4 M $MgCl_2$ Solution to make 5 mM final concentration.
7. Add 2 mL DNAse Solution. Stir until the solution is the viscosity of water.
8. Freeze overnight at –20˚C.
9. Thaw solution in warm water bath.
10. Stir an additional 10 minutes after thawing to allow the DNAse to work again.
11. Spin down in 2 centrifuge bottles at 8000 x g for 20 minutes. Discard supernatant.
12. Re-suspend pellets in 300 mL or more each Triton Solution. Chop briefly, keeping the pellets on ice as much as possible. Spin down at 8000 x g for 20 minutes and discard supernatant.
13. Repeat step 12 three or more times.
14. Re-suspend pellets into 300 mL or more each Tris Solution. Chop briefly, keeping the pellets on ice as much as possible. Spin down at 8000 x g for 20 minutes and discard supernatant.
15. Repeat step 14 two or more times.
16. Re-suspend/dissolve the pellets and chop in ~ 200 mL Urea Solution.
17. Spin down at 20˚C, 15,000 x g for 30 minutes. Filter through a 0.2 μm filter.

***Solutions (**Note: Do not add DTT until just before use!)*:**

Solution Name	Stock	For 1 L, 1x	For 1 L 10x
Sucrose			
50 mM Tris, pH 8.0	1 M Tris, pH 8.0	50 mL	
25% Sucrose	Dry, FW = 342.3	250 g	N/A
1 mM EDTA	0.5 M EDTA	2 mL	
0.1% NaN_3	20 % NaN_3	5 mL	
10 mM DTT	Dry, FW = 154.04	1.54 g	
Deoxycholate-Triton			
1 % Deoxycholic Acid	Dry, FW = 414.6	10 g	100 g
1 % Triton X-100	100% solution	10 mL	100 mL
20 mM Tris, pH 7.5	1 M Tris, pH 7.5	20 mL	200 mL
100 mM NaCl	5 M NaCl	20 mL	200 mL
0.1% NaN_3	20 % NaN_3	5 mL	50 mL
10 mM DTT	Dry, FW = 154.04	1.54 g	N/A
Triton			
0.5 % Triton X-100	100% solution	5 mL	50 mL
50 mM Tris, pH 8.0	1 M Tris, pH 8.0	50 mL	500 mL
100 mM NaCl	5 M NaCl	20 mL	200 mL
1 mM EDTA	0.5 M EDTA	2 mL	20 mL
0.1% NaN_3	20 % NaN_3	5 mL	50 mL
1 mM DTT	Dry, FW = 154.04	0.154 g	N/A
Tris			
50 mM Tris, pH 8.0	1 M Tris, pH 8.0	50 mL	500 mL
1 mM EDTA	0.5 M EDTA	2 mL	20 mL
0.1% NaN_3	20 % NaN_3	5 mL	50 mL
1 mM DTT	Dry, FW = 154.04	0.154 g	N/A
Urea			
8 M Urea	Dry, FW = 60.06	480.48 g	N/A
25 mM MES, pH 6.0	Dry, FW = 213.3	5.33 g	
10 mM EDTA	0.5 M EDTA	20 mL	
0.1 mM DTT	Dry, FW = 154.04	15.4 mg	
DNAse		***Make 50 mL, not 1 L**	
75 mM NaCl	5 M NaCl	0.75 mL	N/A
50% Glycerol	100% Glycerol	25 mL	
2 mg/mL DNAse Sigma Cat# D25	Dry	100 mg	

Modifying Free Cysteines on Proteins Using Maleimides

Protocol modified from Jennifer Stone

Procedure**:**

1. If the protein is not freshly reduced, add DTT to 5 mM and incubate the protein overnight at room temperature.
    a. If the protein has been stored in DTT and is freshly purified by anion exchange, there is no need to re-reduce. Skip to step 2.
2. Either column-purify, dialyze, or concentrate/switch buffers to remove DTT and put protein in reaction buffer (I've tested PBS, Tris and Phosphate pH 6-8. If the pH is too high, the maleimide reagent will also modify primary amines such as lysine residues).
    a. I prefer to use the NAP5 drip columns (GE Healthcare cat# 17-0853-01) because the procedure is quick and easys.
    b. The MHC class I complexes elute between 0.75 mL to 1.5 mL.
3. Immediately check the concentration of the protein, which should be no lower than 10 μM, otherwise hydrolysis of maleimide will occur
4. Add the maleimide reagent to a final concentration so that you get between a 5-10 fold molar excess of maleimide to your freshly reduced protein.
    a. Your maleimide stock solution should be made in DMSO at a concentration greater than 5-10 mM. The reason being is that you want to be able to dilute your DMSO and still have a high enough concentration to prevent hydrolysis. I typically do at least a 1:20 dilution into the protein solution.
5. Mix well and store the solution at room temperature for 1-2 hours. If the probe is fluorescent, store the solution in the dark to prevent fluorescence quenching.
6. After the incubation, you might consider adding DTT to a final concentration of 5 mM to quench any remaining unreacted maleimide, although I've found that this isn't necessary.
7. HPLC purify away the excess maleimide and DTT (if added in step 6) using a Superdex 200 column (GE Healthcare Cat# 17-5175-01) set up on the Biocad.
8. Pool and concentrate the peak fractions.
9. Measure the final concentration and, if applicable, the degree of labeling.

Modifying Primary Amines on Proteins Using NHS Esters
Protocol modified from Pierce and Zu Shen

Procedure**:**

1. Make sure your protein is stored in an amine-free buffer at pH 7.25-8.0. Basically, avoid Tris or glycine buffers. I typically use phosphate or PBS at pH 8.0.
2. Check the concentration of the protein, which should be no lower than 10 μM, otherwise the reaction will not go efficiently.
3. Add the NHS ester reagent to a final concentration of so that you get a 10-20-fold molar excess of NHS ester to your protein. Your NHS ester stock solution should be made using DMSO at a concentration around 4 mM. The reason being is that you want to be able to dilute your DMSO and still have a high enough concentration to prevent hydrolysis of the succinimidyl ester. I typically do at least a 1:20 dilution into the protein solution.
4. Mix well and store the solution at room temperature for 1-2 hours. If the probe is fluorescent, store the solution in the dark to prevent fluorescence quenching.
5. After the incubation, add a 50-fold molar excess of ethanolamine relative to the amount of NHS ester added, to quench any remaining unreacted NHS ester.
    a. Quenching with Tris, followed by multiple rounds of dialysis is insufficient to quench the NHS ester.
6. HPLC purify away the excess NHS ester using a Superdex 200 (GE Healthcare Cat# 17-5175-01) column set up on the biocad.
7. Pool and concentrate the peak fractions.
8. Measure the final concentration and, if applicable, the degree of labeling.

Refolding of MHC-I Complexes
Protocol modified from NIH Tetramer Facility

Procedure**:**

1. Pre-chill 1 L of Folding Buffer in a 1.5 L beaker to 4°C.
2. Add 5 mM reduced glutathione, 0.5 mM oxidized glutathione, and 0.2 mM PMSF to cold Folding Buffer, stirring quickly and try to avoid foaming!
3. Calculate 1 μM heavy chain urea-solubilized subunit and add to 4 mL of Injection Buffer.
4. Calculate 2 μM β_2M urea-solubilized subunit and add to the same 4 mL Injection Buffer.
5. Weigh out 10-20 mg of peptide and dissolve in 0.5 mL DMSO.
6. Add peptide solution dropwise to rapidly stirring Folding Buffer.
7. Forcefully inject heavy chain and β_2M to the stirring reaction through a 26-gauge needle as close to the stir-bar as you can. There may be some precipitation or light foaming.
8. Replace the mixture at 4°C and incubate overnight.
9. The following morning, inject **heavy chain and β_2M** in Injection Buffer into the cold, quickly-stirring mixture as above. Replace at 4°C.
10. In the evening, once again inject **heavy chain and β_2M** in Injection Buffer into the cold, quickly-stirring mixture as above. Replace at 4°C and incubate overnight

Solutions**:**

Folding Buffer (pH 8.0)
400 mM L-Arginine (Sigma A5949)
100 mM Tris, pH 8.0
2 mM EDTA

Injection Buffer (pH 4.2)
3 M Guanidine-HCL
10 mM Sodium Acetate
10 mM EDTA

Other Materials
Urea-solubilized heavy chain and light chain (β2M)
Peptide of choice
Reduced glutathione (Sigma G4705)
Oxidized glutathione (Sigma G4501)
100 mM Phenylmethylsulfonyl fluoride (PMSF - Sigma P7626)

Refolding TCRs
Adaptive from Guoqi Li

Procedure

1. Dilute 2uM α + 2uM β (~4.5mg α, 5.5mg β TCRs in 1.5ml Guanidine buffer with 4uM DTT.
2. Add to 100ml Refolding Buffer:
3. Mix at 4 degrees C for 24 hours.
4. Transfer to dialysis bags.
5. Dialyze in 10-14L dH2O at 4 degrees C for 12 hours.
6. Switch dialysis buffer to 10-14L 10mM Tris pH 8.5 at 4 degrees C for 12 hours.
7. Concentrate with centricon/amicon.

Guanidine Buffer:

- 6M Guanidine HCl
- 10mM Sodium Acetate
- 10mM EDTA

    - For 20ml:
        - Guanidine: 95.66g/mol x 6M x 20ml = 1.91g
        - EDTA: 2ml of 0.5M stock
        - DTT: 30ul of 1M stock

Refolding Buffer

- 100mM Tris pH 8.5
- 400mM L-Arginine
- 5M Urea
- 5mM Cysteine (reduced), dissolve in 1ml 1M NaOH
- 0.5mM Cystine (oxidized), dissolve in 1ml 1M HCl
- 0.2mM PMSF

    - For 100ml:
        - Tris: 2.85 1ml Tris HCL, 7.15ml Tris base
            - OR 10ml of 1M Tris pH 8.5
        - Arginine: 210g/mol x 0.4M x 100ml = 8.4g
        - Urea: 60.06g/mol x 5M x 100ml = 30.03g
        - Cysteine: 121.16g/mol x 5mM x 100ml = 60.5mg
        - Cystine: 240g/mol x 0.5mM x 100ml = 12mg
        - PMSF: 100ul of 0.2mM stock

APPENDIX B

BIBLIOGRAPHY

Achour, a, Persson, K., Harris, R. a, Sundbäck, J., Sentman, C.L., Lindqvist, Y., Schneider, G., and Kärre, K. (1998). The crystal structure of H-2Dd MHC class I complexed with the HIV-1-derived peptide P18-I10 at 2.4 A resolution: implications for T cell and NK cell recognition. Immunity *9*, 199–208.

Adams, J., Narayanan, S., Liu, B., Birnbaum, M., Kruse, A., Bowerman, N., Chen, W., Levin, A., Connolly, J., Zhu, C., et al. (2011). T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. Immunity *35*, 681–693.

Afonine, P. V., Moriarty, N.W., Mustyakimov, M., Sobolev, O. V., Terwilliger, T.C., Turk, D., Urzhumtsev, A., and Adams, P.D. (2015). FEM: Feature-enhanced map. Acta Crystallogr. Sect. D Biol. Crystallogr. *71*, 646–666.

Al-Lazikani, B., Lesk, A.M., and Chothia, C. (2000). Canonical structures for the hypervariable regions of T cell αβ receptors. J. Mol. Biol. *295*, 979–995.

Amria, S., Hajiaghamohseni, L.M., Harbeson, C., Zhao, D., Goldstein, O., Blum, J.S., and Haque, A. (2008). HLA-DM negatively regulates HLA-DR4-restricted collagen pathogenic peptide presentation and T cell recognition. Eur. J. Immunol. *38*, 1961–1970.

Anders, A.K., Call, M.J., Schulze, M.S.E.D., Fowler, K.D., Schubert, D.A., Seth, N.P., Sundberg, E.J., and Wucherpfennig, K.W. (2011). HLA-DM captures partially empty HLA-DR molecules for catalyzed removal of peptide. Nat. Immunol. *12*, 54–61.

Arden, B. (1998). Conserved motifs in T-cell receptor CDR1 and CDR2: Implications for ligand and CD8 co-receptor binding. Curr. Opin. Immunol. *10*, 74–81.

Arden, B., Clark, S.P., Kabelitz, D., and Mak, T.W. (1995). Mouse T-cell receptor variable gene segment families. Immunogenetics *42*, 501–530.

Babbitt, B.P., Allen, P.M., Matsueda, G., Haber, E., and Unanue, E.R. (1985). Binding of immunogenic peptides to Ia histocompatibility molecules. Nature *317*, 359–361.

Baker, B.M., Scott, D.R., Blevins, S.J., and Hawse, W.F. (2012). Structural and dynamic control of T-cell receptor specificity, cross-reactivity, and binding mechanism. Immunol. Rev. *250*, 10–31.

Berg, L.J., Pullen, A.M., de St. Groth, B., Mathis, D., Benoist, C., and Davis, M.M. (1989). Antigen/MHC-specific T cells are preferentially exported from the thymus in the presence of their MHC ligand. Cell *58*, 1035–1046.

Beringer, D.X., Kleijwegt, F.S., Wiede, F., van der Slik, A.R., Loh, K.L., Petersen, J., Dudek, N.L., Duinkerken, G., Laban, S., Joosten, A., et al. (2015). T cell receptor reversed polarity recognition of a self-antigen major histocompatibility complex. Nat. Immunol. *16*, 1153–1161.

Bevan, M.J. (1977). In a radiation chimaera, host H-2 antigens determine immune responsiveness of donor cytotoxic cells. Nature *269*, 417–418.

Birnbaum, M.E., Mendoza, J.L., Sethi, D.K., Dong, S., Glanville, J., Dobbins, J., Davis, M.M., Wucherpfennig, K.W., Garcia, K.C., Özkan, E., et al. (2014). Deconstructing the peptide-MHC specificity of t cell recognition. Cell *157*, 1073–1087.

Bjorkman, P.J., and Parham, P. (1990). Structure, Function, and Diversity of Class I Major Histocompatibility Complex Molecules. Annu. Rev. Biochem. *59*, 253–288.

Blackman, M., Yagüe, J., Kubo, R., Gay, D., Coleclough, C., Palmer, E., Kappler, J., and Marrack, P. (1986). The T cell repertoire may be biased in favor of MHC recognition. Cell *47*, 349–357.

Borbulevych, O.Y., Piepenbrink, K.H., Gloor, B.E., Scott, D.R., Sommese, R.F., Cole, D.K., Sewell, A.K., and Baker, B.M. (2009). T Cell Receptor Cross-reactivity Directed by Antigen-Dependent Tuning of Peptide-MHC Molecular Flexibility. Immunity *31*, 885–896.

Borbulevych, O.Y., Santhanagopolan, S.M., Hossain, M., and Baker, B.M. (2011a). TCRs used in cancer gene therapy cross-react with MART-1/Melan-A tumor antigens via distinct mechanisms. J. Immunol. *187*, 2453–2463.

Borbulevych, O.Y., Santhanagopolan, S.M., Hossain, M., and Baker, B.M. (2011b). TCRs used in cancer gene therapy cross-react with MART-1/Melan-A tumor antigens via distinct mechanisms. J. Immunol. *187*, 2453–2463.

Borbulevych, O.Y., Piepenbrink, K.H., and Baker, B.M. (2011c). Conformational Melding Permits a Conserved Binding Geometry in TCR Recognition of Foreign and Self Molecular Mimics. J. Immunol. *186*, 2950–2958.

Bridgeman, J.S., Sewell, A.K., Miles, J.J., Price, D.A., and Cole, D.K. (2012). Structural and biophysical determinants of αβ T-cell antigen recognition. Immunology *135*, 9–18.

Brown, J.H., Jardetzky, T.S., Gorga, J.C., Stern, L.J., Urban, R.G., Strominger, J.L., and Wiley, D.C. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature *364*, 33–39.

Bulek, A.M., Cole, D.K., Skowera, A., Dolton, G., Gras, S., Madura, F., Fuller, A., Miles, J.J., Gostick, E., Price, D.A., et al. (2012). Structural basis for the killing of human beta cells by CD8+ T cells in type 1 diabetes. Nat. Immunol. *13*, 283–289.

Burmeister, W.P., Gastinel, L.N., Simister, N.E., Blum, M.L., and Bjorkman, P.J. (1994). Crystal structure at 2.2 A resolution of the MHC-related neonatal Fc receptor. Nature *372*, 336–343.

Calis, J.J.A., and Rosenberg, B.R. (2014). Characterizing immune repertoires by high throughput sequencing : strategies and applications. Trends Immunol. *35*, 581–590.

Cameron, B.J., Gerry, A.B., Dukes, J., Harper, J. V, Kannan, V., Bianchi, F.C., Grand, F., Brewer, J.E., Gupta, M., Plesa, G., et al. (2013). Identification of a Titin-Derived HLA-A1 – Presented Peptide as a Cross-Reactive Target for Engineered MAGE A3 – Directed

T Cells. Sci. Transl. Med. *5*, 197ra103-197ra103.

Carlson, C.S., Emerson, R.O., Sherwood, A.M., Desmarais, C., Chung, M.-W., Parsons, J.M., Steen, M.S., LaMadrid-Herrmannsfeldt, M. a, Williamson, D.W., Livingston, R.J., et al. (2013). Using synthetic templates to design an unbiased multiplex PCR assay. Nat. Commun. *4*, 2680.

Che, J.W., Selin, L.K., and Welsh, R.M. (2015). Evaluation of non-reciprocal heterologous immunity between unrelated viruses. Virology *482*, 89–97.

Chen, H.D., Fraire, A.E., Joris, I., Welsh, R.M., and Selin, L.K. (2003). Specific history of heterologous virus infections determines anti-viral immunity and immunopathology in the lung. Am. J. Pathol. *163*, 1341–1355.

Chicz, R.M., Urban, R.G., Lane, W.S., Gorga, J.C., Stern, L.J., Vignali, D.A., and Strominger, J.L. (1992). Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. Nature *358*, 764–768.

Chodon, T., Comin-Anduix, B., Chmielowski, B., Koya, R.C., Wu, Z., Auerbach, M., Ng, C., Avramis, E., Seja, E., Villanueva, A., et al. (2014). Adoptive transfer of MART-1 T-cell receptor transgenic lymphocytes and dendritic cell vaccination in patients with metastatic melanoma. Clin. Cancer Res. *20*, 2457–2465.

Chou, C.-L., and Sadegh-Nasseri, S. (2000). HLA-DM recognizes the flexible conformation of major histocompatibility complex class II. J. Exp. Med. *192*, 1697–1706.

Colf, L.A., Bankovich, A.J., Hanick, N.A., Bowerman, N.A., Jones, L.L., Kranz, D.M.D., and Garcia, K.C. (2007). How a Single T Cell Receptor Recognizes Both Self and Foreign MHC. Cell *129*, 135–146.

Collins, E.J., and Riddle, D.S. (2008). TCR-MHC docking orientation: Natural selection, or thymic selection? Immunol. Res. *41*, 267–294.

Cornberg, M., Sheridan, B.S., Saccoccio, F.M., Brehm, M. a, and Selin, L.K. (2007). Protection against vaccinia virus challenge by CD8 memory T cells resolved by molecular mimicry. J. Virol. *81*, 934–944.

Cornberg, M., Clute, S.C., Watkin, L.B., Saccoccio, F.M., Kim, S.-K.S.-K.S.-K., Naumov, Y.N., Brehm, M.A., Aslan, N., Welsh, R.M., and Selin, L.K. (2010). CD8 T Cell Cross-Reactivity Networks Mediate Heterologous Immunity in Human EBV and Murine Vaccinia Virus Infections. J. Immunol. *184*, 2825–2838.

Correia-Neves, M., Waltzinger, C., Mathis, D., and Benoist, C. (2001). The shaping of the T cell repertoire. Immunity *14*, 21–32.

Corrigan-Curay, J., Kiem, H.P., Baltimore, D., O'Reilly, M., Brentjens, R.J., Cooper, L., Forman, S., Gottschalk, S., Greenberg, P., Junghans, R., et al. (2014). T-cell immunotherapy: Looking forward. Mol. Ther. *22*, 1564–1574.

Crawford, F., Huseby, E., White, J., Marrack, P., and Kappler, J.W. (2004). Mimotopes for alloreactive and conventional T cells in a peptide-MHC display library. PLoS Biol. *2*,

523–533.

Crooks, G.E., Hon, G., Chandonia, J.-M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res. *14*, 1188–1190.

Dai, S., Huseby, E.S., Rubtsova, K., Scott-Browne, J., Crawford, F., Macdonald, W.A., Marrack, P., and Kappler, J.W. (2008). Crossreactive T Cells Spotlight the Germline Rules for αβ T Cell-Receptor Interactions with MHC Molecules. Immunity *28*, 324–334.

Danchin, E., Vitiello, V., Vienne, A., Richard, O., Gouret, P., McDermott, M.F., and Pontarotti, P. (2004). The major histocompatibility complex origin. Immunol Rev *198*, 216–232.

Dash, P., Fiore-Gartland, A.J., Hertz, T., Wang, G.C., Sharma, S., Souquette, A., Crawford, J.C., Clemens, E.B., Nguyen, T.H.O., Kedzierska, K., et al. (2017). Quantifiable predictive features define epitope-specific T cell receptor repertoires. Nature *547*, 89–93.

Davis, M.M., and Bjorkman, P.J. (1988). T-cell antigen receptor genes and T-cell recognition. Nature *334*, 395–402.

Deng, L., Langley, R.J., Wang, Q., Topalian, S.L., and Mariuzza, R.A. (2012). Structural insights into the editing of germ-line-encoded interactions between T-cell receptor and MHC class II by V CDR3. Proc. Natl. Acad. Sci. *109*, 14960–14965.

Ding, Y.H., Baker, B.M., Garboczi, D.N., Biddison, W.E., and Wiley, D.C. (1999). Four A6-TCR/peptide/HLA-A2 structures that generate very different T cell signals are nearly identical. Immunity *11*, 45–56.

Felix, N.J., Donermeyer, D.L., Horvath, S., Walters, J.J., Gross, M.L., Suri, A., and Allen, P.M. (2007). Alloreactive T cells respond specifically to multiple distinct peptide-MHC complexes. Nat. Immunol. *8*, 388–397.

Feng, D., Bond, C.J., Ely, L.K., Maynard, J., and Garcia, K.C. (2007). Structural evidence for a germline-encoded T cell receptor-major histocompatibility complex interaction “codon.” Nat. Immunol. *8*, 975–983.

Ferrante, A., and Gorski, J. (2010). Cutting edge: HLA-DM-mediated peptide exchange functions normally on MHC class II-peptide complexes that have been weakened by elimination of a conserved hydrogen bond. J. Immunol. *184*, 1153–1158.

Fink, P.J., and Bevan, M.J. (1995). Positive Selection of Thymocytes. Adv. Immunol. *59*, 99–134.

Garcia, K.C., Teyton, L., and Wilson, I.A. (1999). Structural basis of T cell recognition. Annu. Rev. Immunol. *17*, 369–397.

Garcia, K.C., Adams, J.J., Feng, D., Ely, L.K., Christopher Garcia, K., Adams, J.J., Feng, D., Ely, L.K., Garcia, K.C., Adams, J.J., et al. (2009a). The molecular basis of TCR germline bias for MHC is surprisingly simple. Nat. Immunol. *10*, 143–147.

Garcia, K.C., Adams, J.J., Feng, D., and Ely, L.K. (2009b). The molecular basis of TCR germline bias for MHC is surprisingly simple. Nat. Immunol. *10*, 143–147.

Garman, R.D., Ko, J.-L., Vulpe, C.D., and Raulet, D.H. (1986). T-cell receptor variable region gene usage in T-cell populations. Proc. Natl. Acad. Sci. U. S. A. *83*, 3987–3991.

Glanville, J., Huang, H., Nau, A., Hatton, O., Wagar, L.E., Rubelt, F., Ji, X., Han, A., Krams, S.M., Pettus, C., et al. (2017). Identifying specificity groups in the T cell receptor repertoire. Nature *547*, 94–98.

Gundlach, B.R., Wiesmiiller, K., Junt, T., Kienle, S., Jung, G., and Walden, P. (1996). Determination of T cell epitopes with random peptide libraries. Peptides.

Günther, S., Schlundt, A., Sticht, J., Roske, Y., Heinemann, U., Wiesmüller, K.-H., Jung, G., Falk, K., Rötzschke, O., and Freund, C. (2010). Bidirectional binding of invariant chain peptides to an MHC class II molecule. Proc. Natl. Acad. Sci. U. S. A. *107*, 22219–22224.

Hammer, J., Valsasnini, P., Tolba, K., Bolin, D., Higelin, J., Takacs, B., and Sinigaglia, F. (1993). Promiscuous and allele-specific anchors in HLA-DR-binding peptides. Cell *74*, 197–203.

Hanvesakul, R., Maillere, B., Briggs, D., Baker, R., Larché, M., and Ball, S. (2007). Indirect recognition of T-cell epitopes derived from the α3 and transmembrane domain of HLA-A2. Am. J. Transplant. *7*, 1148–1157.

Hartman, I.Z., Kim, A., Cotter, R.J., Walter, K., Dalai, S.K., Boronina, T., Griffith, W., Lanar, D.E., Schwenk, R., Krzych, U., et al. (2010). A reductionist cell-free major histocompatibility complex class II antigen processing system identifies immunodominant epitopes. Nat. Med. *16*, 1333–1340.

Hu, G., Modrek, B., Riise Stensland, H.M.F., Saarela, J., Pajukanta, P., Kustanovich, V., Peltonen, L., Nelson, S.F., and Lee, C. (2002). Efficient discovery of single-nucleotide polymorphisms in coding regions of human genes. Pharmacogenomics J. *2*, 236–242.

Huseby, E.S., White, J., Crawford, F., Vass, T., Becker, D., Pinilla, C., Marrack, P., and Kappler, J.W. (2005a). How the T cell repertoire becomes peptide and MHC specific. Cell *122*, 247–260.

Huseby, E.S., White, J., Crawford, F., Vass, T., Becker, D., Pinilla, C., Marrack, P., and Kappler, J.W. (2005b). How the T cell repertoire becomes peptide and MHC specific. Cell *122*, 247–260.

Huseby, E.S., Crawford, F., White, J., Marrack, P., and Kappler, J.W. (2006). Interface-disrupting amino acids establish specificity between T cell receptors and complexes of major histocompatibility complex and peptide. Nat. Immunol. *7*, 1191–1199.

Illing, P.T., Vivian, J.P., Dudek, N.L., Kostenko, L., Chen, Z., Bharadwaj, M., Miles, J.J., Kjer-Nielsen, L., Gras, S., Williamson, N. a., et al. (2012). Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. Nature *486*, 1–7.

Izraelson, M., Nakonechnaya, T.O., Moltedo, B., Egorov, E.S., Kasatskaya, S.A., Putintseva, E. V., Mamedov, I.Z., Staroverov, D.B., Shemiakina, I.I., Zakharova, M.Y., et al. (2018). Comparative analysis of murine T-cell receptor repertoires. Immunology *153*, 133–144.

Jameson, S.C., and Bevan, M.J. (1992). Dissection of major histocompatibility complex (MHC) and T cell receptor contact residues in a Kb-restricted ovalbumin peptide and an assessment of the predictive power of MHC-binding motifs. Eur. J. Immunol. *22*, 2663–2667.

Jameson, S.C., Kaye, J., and Gascoigne, N.R.J. (1990). A T cell receptor Vα region selectively expressed in CD4+ cells. J. Immunol. *145*, 1324–1331.

Joffre, O.P., Segura, E., Savina, A., and Amigorena, S. (2012). Cross-presentation by dendritic cells. Nat. Rev. Immunol. *12*, 557–569.

Jones, L.L., Colf, L.A., Stone, J.D., Garcia, K.C., and Kranz, D.M. (2008). Distinct CDR3 conformations in TCRs determine the level of cross-reactivity for diverse antigens, but not the docking orientation. J. Immunol. *181*, 6255–6264.

Jorgensen, J.L., Esser, U., de St. Groth, B., Reay, P.A., and Davis, M.M. (1992). Mapping T-cell receptor-peptide contacts by variant peptide immunization of single-chain transgenics. Nature *355*, 224–230.

Kappler, J.W., Roehm, N., and Marrack, P. (1987). T cell tolerance by clonal elimination in the thymus. Cell *49*, 273–280.

Kim, S.-K., Cornberg, M., Wang, X.Z., Chen, H.D., Selin, L.K., and Welsh, R.M. (2005). Private specificities of CD8 T cell responses control patterns of heterologous immunity. J. Exp. Med. *201*, 523–533.

Kisielow, P., and Von Boehmer, H. (1995). Development and Selection of T Cells: Facts and Puzzles. Adv. Immunol. *58*, 87–209.

Kitamura, H., Iwakabe, K., Yahata, T., Nishimura, S., Ohta, A., Ohmi, Y., Sato, M., Takeda, K., Okumura, K., Van Kaer, L., et al. (1997). CD1d-restricted and TCR-mediated activation of Vα14 NKT cells by glycosylceramides. Science (80-.). *278*, 1626–1629.

Kjer-Nielsen, L., Clements, C.S., Purcell, A.W., Brooks, A.G., Whisstock, J.C., Burrows, S.R., McCluskey, J., and Rossjohn, J. (2003). A Structural Basis for the Selection of Dominant αβ T Cell Receptors in Antiviral Immunity. Immunity *18*, 53–64.

Klein, J. (1986). Natural history of the major histocompatibility complex (New York: Wiley).

Klein, L., Kyewski, B., Allen, P.M., and Hogquist, K.A. (2014). Positive and negative selection of the T cell repertoire: What thymocytes see (and don't see). Nat. Rev. Immunol. *14*, 377–391.

Košmrlj, A., Read, E.L., Qi, Y., Allen, T.M., Altfeld, M., Deeks, S.G., Pereyra, F., Carrington, M., Walker, B.D., and Chakraborty, A.K. (2010). Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. Nature *465*, 350–354.

Kremer, A.N., Van Der Meijden, E.D., Honders, M.W., Goeman, J.J., Wiertz, E.J.H.J., Falkenburg, J.H.F., and Griffioen, M. (2012). Endogenous HLA class II epitopes that are

immunogenic in vivo show distinct behavior toward HLA-DM and its natural inhibitor HLA-DO. Blood *120*, 3246–3255.

Van Laethem, F., Tikhonova, A.N., and Singer, A. (2012). MHC restriction is imposed on a diverse T cell receptor repertoire by CD4 and CD8 co-receptors during thymic selection. Trends Immunol. *33*, 437–441.

Laydon, D.J., Bangham, C.R.M., and Asquith, B. (2015). Estimating T-cell repertoire diversity: limitations of classical estimators and a new approach. Phil Trans R Soc B *370*, 20140291-.

Lazarski, C.A., Chaves, F.A., Jenks, S.A., Wu, S., Richards, K.A., Weaver, J.M., and Sant, A.J. (2005). The kinetic stability of MHC class II:Peptide complexes is a key parameter that dictates immunodominance. Immunity *23*, 29–40.

Lazarski, C.A., Chaves, F.A., and Sant, A.J. (2006). The impact of DM on MHC class II-restricted antigen presentation can be altered by manipulation of MHC-peptide kinetic stability. J. Exp. Med. *203*, 1319–1328.

Li, W., Joshi, M.D., Singhania, S., Ramsey, K.H., and Murthy, A.K. (2014). Peptide Vaccine: Progress and Challenges. Vaccines *2*, 515–536.

Li, Y., Moysey, R., Molloy, P.E., Vuidepot, A.-L., Mahon, T., Baston, E., Dunn, S., Liddy, N., Jacob, J., Jakobsen, B.K., et al. (2005). Directed evolution of human T-cell receptors with picomolar affinities by phage display. Nat. Biotechnol. *23*, 349–354.

Linette, G.P., Stadtmauer, E.A., Maus, M. V, Rapoport, A.P., Levine, B.L., Emery, L., Litzky, L., Bagg, A., Carreno, B.M., Cimino, P.J., et al. (2013). Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. Blood *122*, 863–871.

Liu, Y.C., Miles, J.J., Neller, M.A., Gostick, E., Price, D.A., Purcell, A.W., McCluskey, J., Burrows, S.R., Rossjohn, J., and Gras, S. (2013). Highly divergent T-cell receptor binding modes underlie specific recognition of a bulged viral peptide bound to a human leukocyte antigen class I molecule. J. Biol. Chem. *288*, 15442–15454.

López-Sagaseta, J., Sibener, L. V, Kung, J.E., Gumperz, J., and Adams, E.J. (2012). Lysophospholipid presentation by CD1d and recognition by a human Natural Killer T-cell receptor. EMBO J. *31*, 2047–2059.

Lovitch, S.B., Petzold, S.J., and Unanue, E.R. (2003). Cutting edge: H-2DM is responsible for the large differences in presentation among peptides selected by I-Ak during antigen processing. J. Immunol. *171*, 2183–2186.

Macdonald, W.A., Chen, Z., Gras, S., Archbold, J.K., Tynan, F.E., Clements, C.S., Bharadwaj, M., Kjer-Nielsen, L., Saunders, P.M., Wilce, M.C.J.J., et al. (2009). T Cell Allorecognition via Molecular Mimicry. Immunity *31*, 897–908.

Marrack, P., Scott-Browne, J.P., Dai, S., Gapin, L., and Kappler, J.W. (2008). Evolutionarily conserved amino acids that control TCR-MHC interaction. Annu. Rev. Immunol. *26*, 171–203.

Mason, D. (1998). A very high level of crossreactivity is an essential feature of the T-cell receptor. Immunol. Today *19*, 395–404.

Mathis, D., and Benoist, C. (2004). Back to central tolerance. Immunity *20*, 509–516.

Mathurin, K.S., Martens, G.W., Kornfeld, H., and Welsh, R.M. (2009). CD4 T-Cell-Mediated Heterologous Immunity between Mycobacteria and Poxviruses. J. Virol. *83*, 3528–3539.

Matsumura, M., Fremont, D.H., Peterson, P.A., and Wilson, I.A. (1992). Emerging principles for the recognition of peptide antigens by MHC class I molecules. Science *257*, 927–934.

Matzaraki, V., Kumar, V., Wijmenga, C., and Zhernakova, A. (2017). The MHC locus and genetic susceptibility to autoimmune and infectious diseases.

Mazza, C., Auphan-Anezin, N., Gregoire, C., Guimezanes, A., Kellenberger, C., Roussel, A., Kearney, A., van der Merwe, P.A., Schmitt-Verhulst, A.-M., and Malissen, B. (2007a). How much can a T-cell antigen receptor adapt to structurally distinct antigenic peptides? EMBO J. *26*, 1972–1983.

Mazza, C., Auphan-Anezin, N., Gregoire, C., Guimezanes, A., Kellenberger, C., Roussel, A., Kearney, A., Van Der Merwe, P.A., Schmitt-Verhulst, A.-M., and Malissen, B. (2007b). How much can a T-cell antigen receptor adapt to structurally distinct antigenic peptides? EMBO J. *26*, 1972–1983.

Merkenschlager, M., Graf, D., Lovatt, M., Bommhardt, U., Zamoyska, R., and Fisher, A.G. (1997). How many thymocytes audition for selection? J. Exp. Med. *186*, 1149–1158.

Messaoudi, I., Guevara Patiño, J.A., Dyall, R., LeMaoult, J., and Nikolich-Žugich, J. (2002). Direct link between mhc polymorphism, T cell avidity, and diversity in immune defense. Science (80-.). *298*, 1797–1800.

Mitaksov, V., and Fremont, D.H. (2006). Structural definition of the H-2Kd peptide-binding motif. J. Biol. Chem. *281*, 10618–10625.

Morgan, R.A., Chinnasamy, N., Abate-daga, D.D., Gros, A., Robbins, F., Zheng, Z., Feldman, S.A., Yang, J.C., Sherry, R.M., Phan, Q., et al. (2014). Cancer regressions and neurologic toxicity following anti-MAGE-A3 TCR gene therapy. J Immunother. *36*, 133–151.

Morris, G.P., Ni, P.P., and Allen, P.M. (2011). Alloreactivity is limited by the endogenous peptide repertoire. Proc. Natl. Acad. Sci. U. S. A. *108*, 3695–3700.

Van Der Most, R.G., Murali-Krishna, K., Lanier, J.G., Wherry, E.J., Puglielli, M.T., Blattman, J.N., Sette, A., and Ahmed, R. (2003). Changing immunodominance patterns in antiviral CD8 T-cell responses after loss of epitope presentation or chronic antigenic stimulation. Virology *315*, 93–102.

Murthy, V.L., and Stern, L.J. (1997). The class II MHC protein HLA-DR1 in complex with an endogenous peptide: implications for the structural basis of the specificity of

peptide binding. Structure *5*, 1385–1396.

Narayan, K., Chou, C.-L., Kim, A., Hartman, I.Z., Dalai, S., Khoruzhenko, S., and Sadegh-Nasseri, S. (2007). HLA-DM targets the hydrogen bond between the histidine at position β81 and peptide to dissociate HLA-DR-peptide complexes. Nat. Immunol. *8*, 92–100.

Narayan, K., Su, K.W., Chou, C.-L., Khoruzhenko, S., and Sadegh-Nasseri, S. (2009). HLA-DM mediates peptide exchange by interacting transiently and repeatedly with HLA-DR1. Mol. Immunol. *46*, 3157–3162.

Natarajan, S.K., Stern, L.J., and Sadegh-Nasseri, S. (1999). Sodium dodecyl sulfate stability of HLA-DR1 complexes correlates with burial of hydrophobic residues in pocket 1. J. Immunol. *162*, 3463–3470.

Neefjes, J., M Jongsma, M.L., and Paul, P. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat. Rev. Immunol. *11*, 823–836.

Nussinov, R., and Tsai, C.-J. (2013). Allostery in disease and in drug discovery. Cell *153*, 293–305.

Ostrov, D. a, Grant, B.J., Pompeu, Y. a, Sidney, J., Harndahl, M., Southwood, S., Oseroff, C., Lu, S., Jakoncic, J., de Oliveira, C.A.F., et al. (2012). Drug hypersensitivity caused by alteration of the MHC-presented self-peptide repertoire. Proc. Natl. Acad. Sci. U. S. A.

Painter, C.A., and Stern, L.J. (2012). Conformational variation in structures of classical and non-classical MHCII proteins and functional implications. Immunol. Rev. *250*, 144–157.

Painter, C.A., Negroni, M.P., Kellersberger, K.A., Zavala-Ruiz, Z., Evans, J.E., and Stern, L.J. (2011). Conformational lability in the class II MHC 310 helix and adjacent extended strand dictate HLA-DM susceptibility and peptide exchange. Proc. Natl. Acad. Sci. *108*, 19329–19334.

Pereyra, F., Jia, X., McLaren, P.J., Telenti, A., De Bakker, P.I.W., and Walker, B.D. (2010). The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science (80-.). *330*, 1551–1557.

Pos, W., Sethi, D.K., Call, M.J., Schulze, M.S.E.D., Anders, A.K., Pyrdol, J., and Wucherpfennig, K.W. (2012). Crystal structure of the HLA-DM-HLA-DR1 complex defines mechanisms for rapid peptide selection. Cell *151*, 1557–1568.

Pu, Z., Carrero, J.A., and Unanue, E.R. (2002). Distinct recognition by two subsets of T cells of an MHC class II-peptide complex. Proc. Natl. Acad. Sci. U. S. A. *99*, 8844–8849.

Rammensee, H.G. (1995). Chemistry of peptides associated with MHC class I and class II molecules. Curr. Opin. Immunol. *7*, 85–96.

Rammensee, H.G., Falk, K., and Rötzschke, O. (1993). Peptides naturally presented by MHC class I molecules. Annu. Rev. Immunol. *11*, 213–244.

Rapoport, A.P., Stadtmauer, E.A., Binder-Scholl, G.K., Goloubeva, O., Vogl, D.T.,

Lacey, S.F., Badros, A.Z., Garfall, A., Weiss, B., Finklestein, J., et al. (2015). NY-ESO-1-specific TCR-engineered T cells mediate sustained antigen-specific antitumor effects in myeloma. Nat. Med. *21*, 914–921.

Reiser, J.B., Darnault, C., Grégoire, C., Mosser, T., Mazza, G., Kearney, A., van der Merwe, P.A., Fontecilla-Camps, J.C., Housset, D., and Malissen, B. (2003). CDR3 loop flexibility contributes to the degeneracy of TCR recognition. Nat. Immunol. *4*, 241–247.

Robbins, P.F., Kassim, S.H., Tran, T.L.N., Crystal, J.S., Morgan, R.A., Feldman, S.A., Yang, J.C., Dudley, M.E., Wunderlich, J.R., Sherry, R.M., et al. (2015). A pilot trial using lymphocytes genetically engineered with an NY-ESO-1-reactive T-cell receptor: Long-term follow-up and correlates with response. Clin. Cancer Res. *21*, 1019–1027.

Robins, H.S., Campregher, P. V, Srivastava, S.K., Wacher, A., Turtle, C.J., Kahsai, O., Riddell, S.R., Warren, E.H., and Carlson, C.S. (2010). Comprehensive assessment of T-cell receptor β-chain diversity in αβ T cells. Blood *114*, 4099–4107.

Roche, P.A., and Cresswell, P. (1990). High-affinity binding of an influenza hemagglutinin-derived peptide to purified HLA-DR. J. Immunol. *144*, 1849–1856.

Rosati, E., Dowds, C.M., Liaskou, E., Henriksen, E.K.K., Karlsen, T.H., and Franke, A. (2017). Overview of methodologies for T-cell receptor repertoire analysis. BMC Biotechnol. *17*, 1–16.

Rossjohn, J., Gras, S., Miles, J.J., Turner, S.J., Godfrey, D.I., and McCluskey, J. (2015). T cell antigen receptor recognition of antigen-presenting molecules. Annu. Rev. Immunol. *33*, 169–200.

Rubtsova, K., Scott-Browne, J.P., Crawford, F., Dai, S., Marrack, P., and Kappler, J.W. (2009). Many different Vbeta CDR3s can reveal the inherent MHC reactivity of germline-encoded TCR V regions. Proc. Natl. Acad. Sci. U. S. A. *106*, 7951–7956.

Rudolph, M.G., Stanfield, R.L., and Wilson, I.A. (2006a). How Tcrs Bind Mhcs, Peptides, and Coreceptors. Annu. Rev. Immunol. *24*, 419–466.

Rudolph, M.G., Stanfield, R.L., and Wilson, I.A. (2006b). How Tcrs Bind Mhcs, Peptides, and Coreceptors. Annu. Rev. Immunol. *24*, 419–466.

Saito, Y., Peterson, P.A., and Matsumura, M. (1993). Quantitation of peptide anchor residue contributions to class I major histocompatibility complex molecule binding. J. Biol. Chem.

Sánchez, L.M. (1999). Crystal structure of human ZAG, a fat-depleting factor related to MHC molecules. Science (80-.). *283*, 1914–1919.

Sánchez, L.M., López-Otín, C., and Bjorkman, P.J. (1997). Biochemical characterization and crystalization of human Zn-alpha2-glycoprotein, a soluble class I major histocompatibility complex homolog. Proc. Natl. Acad. Sci. U. S. A. *94*, 4626–4630.

Sant, A.J., Chaves, F.A., Jenks, S.A., Richards, K.A., Menges, P., Weaver, J.M., and Lazarski, C.A. (2005). The relationship between immunodominance, DM editing, and the kinetic stability of MHC class II:peptide complexes. Immunol. Rev. *207*, 261–278.

Sato, A.K., Zarutskie, J.A., Rushe, M.M., Lomakin, A., Natarajan, S.K., Sadegh-Nasseri, S., Benedek, G.B., and Stern, L.J. (2000). Determinants of the peptide-induced conformational change in the human class II major histocompatibility complex protein HLA-DR1. J Biol Chem *275*, 2165–2173.

Sawicka, M., Stritesky, G.L., Reynolds, J., Abourashchi, N., Lythe, G., Molina-París, C., and Hogquist, K.A. (2014). From pre-DP, post-DP, SP4, and SP8 thymocyte cell counts to a dynamical model of cortical and medullary selection. Front. Immunol. *5*, 1–14.

Schlundt, A., Günther, S., Sticht, J., Wieczorek, M., Roske, Y., Heinemann, U., and Freund, C. (2012). Peptide linkage to the α-subunit of MHCII creates a stably inverted antigen presentation complex. J. Mol. Biol. *423*, 294–302.

Schulze, M.S.E.D., Anders, A.K., Sethi, D.K., and Call, M.J. (2013). Disruption of Hydrogen Bonds between Major Histocompatibility Complex Class II and the Peptide N-Terminus Is Not Sufficient to Form a Human Leukocyte Antigen-DM Receptive State of Major Histocompatibility Complex Class II. PLoS One *8*, 1–15.

Scott, B., Blüthmann, H., Teh, H.S., and Von Boehmer, H. (1989). The generation of mature T cells requires interaction of the αβ T-cell receptor with major histocompatibility antigens. Nature *338*, 591–593.

Seedhom, M.O., Jellison, E.R., Daniels, K.A., and Welsh, R.M. (2009). High Frequencies of Virus-Specific CD8+ T-Cell Precursors. J. Virol. *83*, 12907–12916.

Selin, L.K., Nahill, S.R., and Welsh, R.M. (1994). Cross-reactivities in memory cytotoxic T lymphocyte recognition of heterologous viruses. J. Exp. Med. *179*, 1933–1943.

Selin, L.K., Varga, S.M., Wong, I.C., and Welsh, R.M. (1998). Protective heterologous antiviral immunity and enhanced immunopathogenesis mediated by memory T cell populations. J. Exp. Med. *188*, 1705–1715.

Sewell, A.K. (2012). Why must T cells be cross-reactive ? Nat. Publ. Gr. *12*, 669–677.

Sharfstein, S.T. (2018). Non-protein biologic therapeutics. Curr. Opin. Biotechnol. *53*, 65–75.

Shen, Z.T., Brehm, M.A., Daniels, K.A., Sigalov, A.B., Selin, L.K., Welsh, R.M., and Stern, L.J. (2010). Bi-specific MHC heterodimers for characterization of cross-reactive T cells. J. Biol. Chem. *285*, 33144–33153.

Shen, Z.T., Nguyen, T.T., Daniels, K.A., Welsh, R.M., and Stern, L.J. (2013). Disparate epitopes mediating protective heterologous immunity to unrelated viruses share peptide-MHC structural features recognized by cross-reactive T cells. J. Immunol. *191*, 5139–5152.

Sim, B.-C., Zerva, L., Greene, M.I., and Gascoigne, N.R.J. (1996). Control of MHC restriction by TCR V(α) CDR1 and CDR2. Science (80-.). *273*, 963–966.

Smith, H.J., Hanvesakul, R., Bentall, A., Shabir, S., Morgan, M.D., Briggs, D., Cockwell, P., Borrows, R., Larché, M., and Ball, S. (2011). T lymphocyte responses to nonpolymorphic HLA-derived peptides are associated with chronic renal allograft

dysfunction. Transplantation *91*, 279–286.

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Front. Zool. *2*, 1–18.

Stadinski, B.D., Trenh, P., Smith, R.L., Bautista, B., Huseby, P.G., Li, G., Stern, L.J., and Huseby, E.S. (2011). A role for differential variable gene pairing in creating T cell receptors specific for unique major histocompatibility ligands. Immunity *35*, 694–704.

Stern, L.J., and Wiley, D.C. (1994). Antigenic peptide binding by class I and class II histocompatibility proteins. Structure *2*, 245–251.

Stern, L.J., Brown, J.H., Jardetzky, T.S., Gorga, J.C., Urban, R.G., Strominger, J.L., and Wiley, D.C. (1994). Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. Nature *368*, 215–221.

Stern, L.J., Potolicchio, I., and Santambrogio, L. (2006). MHC class II compartment subtypes: Structure and function. Curr. Opin. Immunol. *18*, 64–69.

Stratikos, E., Wiley, D.C., and Stern, L.J. (2004). Enhanced Catalytic Action of HLA-DM on the Exchange of Peptides Lacking Backbone Hydrogen Bonds between their N-Terminal Region and the MHC Class II α-Chain. J. Immunol. *172*, 1109–1117.

Stritesky, G.L., Xing, Y., Erickson, J.R., Kalekar, L.A., Wang, X., Mueller, D.L., Jameson, S.C., and Hogquist, K.A. (2013). Murine thymic selection quantified using a unique method to capture deleted T cells. Proc. Natl. Acad. Sci. *110*, 4679–4684.

Tikhonova, A.N., Van Laethem, F., Hanada, K.I., Lu, J., Pobezinsky, L.A., Hong, C., Guinter, T.I., Jeurling, S.K., Bernhardt, G., Park, J.H., et al. (2012). αβ T Cell Receptors that Do Not Undergo Major Histocompatibility Complex-Specific Thymic Selection Possess Antibody-like Recognition Specificities. Immunity *36*, 79–91.

Treiner, E., Duban, L., Bahram, S., Radosavljevic, M., Wanner, V., Tilloy, F., Affaticati, P., Gilfillan, S., and Lantz, O. (2003). Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. Nature *422*, 164–169.

Trowsdale, J., and Knight, J.C. (2013). Major Histocompatibility Complex Genomics and Human Disease. Annu. Rev. Genomics Hum. Genet. *14*, 301–323.

Turtle, C.J., Hanafi, L., Berger, C., Gooley, T.A., Cherian, S., Hudecek, M., Sommermeyer, D., Melville, K., Pender, B., Budiarto, T.M., et al. (2016). CD19 CAR – T cells of defined CD4 + : CD8 + composition in adult B cell ALL patients. J Clin Invest *1*, 1–16.

Tynan, F.E., Burrows, S.R., Buckle, A.M., Clements, C.S., Borg, N. a, Miles, J.J., Beddoe, T., Whisstock, J.C., Wilce, M.C., Silins, S.L., et al. (2005). T cell receptor recognition of a “super-bulged” major histocompatibility complex class I-bound peptide. Nat. Immunol. *6*, 1114–1122.

Tynan, F.E., Reid, H.H., Kjer-Nielsen, L., Miles, J.J., Wilce, M.C.J., Kostenko, L., Borg, N. a, Williamson, N. a, Beddoe, T., Purcell, A.W., et al. (2007). A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I

molecule. Nat. Immunol. *8*, 268–276.

Vacic, V., Iakoucheva, L.M., and Radivojac, P. (2006). Two Sample Logo: A graphical representation of the differences between two sets of sequence alignments. Bioinformatics *22*, 1536–1537.

Valkenburg, S.A., Day, E.B., Swan, N.G., Croom, H.A., Carbone, F.R., Doherty, P.C., Turner, S.J., and Kedzierska, K. (2010). Fixing an irrelevant TCRα chain reveals the importance of TCRβ diversity for optimal TCRαβ pairing and function of virus-specific CD8+ T cells. Eur. J. Immunol. *40*, 2470–2481.

Vyas, J.M., Van Der Veen, A.G., and Ploegh, H.L. (2008). The known unknowns of antigen processing and presentation. Nat. Rev. Immunol. *8*, 607–618.

Walzl, G., Tafuro, S., Moss, P., Openshaw, P.J.M., and Hussell, T. (2000). Influenza Virus Lung Infection Protects from Respiratory Syncytial Virus–Induced Immunopathology. J. Exp. Med. *192*, 1317–1326.

Welsh, R.M., Che, J.W., and Brehm, M.A. (2010). Heterologous immunity between viruses. *235*, 244–266.

Wilson, D.H.D.B., Pinilla, C., Wilson, D.H.D.B., Boggiano, C., Judkowski, V., Hemmer, B., Martin, R., Richard, A., Kaye, J., and Houghten, R.A. (2012). Cell Responses to Native Peptide Ligands 1. Cultures.

Wölpl, A., Halder, T., Kalbacher, H., Neumeyer, H., Siemoneit, K., Goldmann, S.F., and Eiermann, T.H. (1998). Human monoclonal antibody with T-cell-like specificity recognizes MHC class I self-peptide presented by HLA-DR1 on activated cells. Tissue Antigens *51*, 258–269.

Wooldridge, L., Ekeruche-Makinde, J., Berg, H.A. Van Den, Skowera, A., Miles, J.J., Tan, M.P., Dolton, G., Clement, M., Llewellyn-lacey, S., Price, D.A., et al. (2012). A single autoimmune T cell receptor recognizes more than a million different peptides. J. Biol. Chem. *287*, 1168–1177.

Wucherpfennig, K.W., Allen, P.M., Celada, F., Cohen, I.R., De Boer, R., Garcia, K.C., Goldstein, B., Greenspan, R., Hafler, D., Hodgkin, P., et al. (2007). Polyspecificity of T cell and B cell receptor recognition. Semin. Immunol. *19*, 216–224.

Yin, Y., and Mariuzza, R. a (2009). The multiple mechanisms of T cell receptor cross-reactivity. Immunity *31*, 849–851.

Yin, L., Calvo-Calle, J.M., Dominguez-Amorocho, O., and Stern, L.J. (2012). HLA-DM Constrains Epitope Selection in the Human CD4 T Cell Response to Vaccinia Virus by Favoring the Presentation of Peptides with Longer HLA-DM-Mediated Half-Lives. J. Immunol. *189*, 3983–3994.

Young, A.C., Zhang, W., Sacchettini, J.C., and Nathenson, S.G. (1994). The three-dimensional structure of H-2Db at 2.4 A resolution: implications for antigen-determinant selection. Cell *76*, 39–50.

Zerrahn, J., Held, W., and Raulet, D.H. (1997). The MHC reactivity of the T cell

repertoire prior to positive and negative selection. Cell *88*, 627–636.

Zhang, C., Anderson, A., and DeLisi, C. (1998). Structural principles that govern the peptide-binding motifs of class I MHC molecules. J. Mol. Biol. *281*, 929–947.

Zhou, Z., Callaway, K.A., Weber, D.A., and Jensen, P.E. (2009). Cutting edge: HLA-DM functions through a mechanism that does not require specific conserved hydrogen bonds in class II MHC-peptide complexes. J. Immunol. *183*, 4187–4191.

Zinkernagel, R.M., and Doherty, P.C. (1974). Restriction of in vitro T cell mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. Nature *248*, 701–702.