The origin and function of a large naïve CD4\(^+\) T cell population

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Ryan William Nelson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Marc K. Jenkins, Ph.D., Advisor

June 2014
Acknowledgements

I would like to thank my family for their unwavering support and encouragement over the years. I thank Jennifer for her assistance, and patience, throughout my time spent in the Jenkins lab. I thank Jim, James, Marion, Kathy, Hamlet, Antonio, Jon, Justin, Miriam, Noah, Dileepan, Mike, Dmitri, Deepali, and Jessica for all of the helpful discussions and afternoons at Stub & Herb’s. To Rebecca and Alexis, thank you for your many efforts and technical support. I thank my past research mentors: Frank Prendergast and Enrico Di Cera for giving me the opportunity to work in their laboratories and encouraging my pursuit of a research career. Many thanks to Yoji Shimizu, Tucker LeBien, Pete Bitterman, Susan Shurson, Nick Berg, and all members of the MD/PhD program for their support and guidance throughout my time in the program. Tim Schacker and Bryce Binstadt, thank you for supporting my continued medical training and for serving as role models of physician scientists. I thank Dan Kaplan, Dan Mueller, and Steve Jameson for forming my thesis committee, providing insights and advice, and for treating me like a colleague. To all of my friends and colleagues at the Center for Immunology, thanks for everything. Most importantly, thanks to my mentor, Marc Jenkins. For making me a better scientist, instilling confidence, having fun, never letting me pay for a beer, and teaching me “nothing is free,” I am in debt to you.
Dedication

To Nancy.
Abstract

The primary goal of immunology is to maximize the immune system’s potential to control or eliminate harmful infection. This can be achieved through vaccination, though current vaccine strategies have had limited success when targeting CD4⁺ T cell-dependent cellular immunity to intracellular infections. CD4⁺ T cells are cells of the adaptive immune system that use clonally-distributed T cell antigen receptors (TCRs) to recognize 9 amino acid peptide antigens bound to Major Histocompatibility Complex II (MHCII) molecules on host cells. Each vertebrate animal contains a large set of CD4⁺ T cells, each with a unique TCR and generated before infection. Here, we studied the repertoire of CD4⁺ T cells in mice to better understand why some non-mouse (foreign) peptide:MHCII ligands (p:MHCII) stimulate stronger immune response than others. We found that naïve CD4⁺ T cell populations specific for different foreign p:MHCII vary considerably in size and that large naïve populations produce more effector cells during an immune response than small ones. Because the T cell repertoire in a mouse is generated by removal in the thymus of clones with TCRs that recognize mouse p:MHCII, we tested the possibility that the size of a given naïve foreign peptide-specific T cell population is shaped by negative selection on similar mouse (self) peptides. We found that identity at 4 amino acid positions was all that was required for two MHCII-binding nonamer peptides to bind the same TCR. This basic principle allowed in silico approaches to reveal that the number of cells in a foreign peptide-specific T cell population was inversely related to the number of MHCII-binding self peptides with the same 4 TCR contact amino acids. Therefore, the size of a given naïve foreign p:MHCII-
specific T cell population is shaped by negative selection due to TCR cross-reaction on similar self p:MHCII ligands.

We next explored the effect of a large foreign p:MHCII-specific T cell population on immunity. Because CD4$^{+}$ T cells protect hosts from infections that have evolved to persist in the phagosomes of infected cells, it was of interest to test this issue during a persistent Salmonella enterica serovar Typhimurium (ST) infection. Remarkably, I found that this one large foreign p:MHCII-specific T cell population played a large role in controlling ST infection, despite the likely presentation of many other ST-derived p:MHCII. This control was associated with numerical and functional stability of the large foreign p:MHCII-specific T cell population for greater than a year after oral infection. This stability was associated with peptide:MHCII-driven proliferation by a small number of T cells in the secondary lymphoid organs that harbored bacteria. Thus, my work could instruct novel subunit vaccine strategies to target large naïve CD4$^{+}$ T cell populations and maintain their responses with longer-lasting peptide delivery.
Table of Contents

Acknowledgements................................................................. i
Dedication............................................................................... ii
Abstract................................................................................ iii
Table of Contents................................................................. v
List of Tables........................................................................ viii
List of Figures......................................................................... ix

Chapter 1: Introduction.......................................................... 1
  Background and Significance
    Adaptive immunity and the Clonal Selection Theory
    Generation and selection of TCR diversity during thymic development
    p:MHCII recognition by CD4⁺ T cells
    The pre-immune T cell repertoire
    Immunodominance
    Statement of thesis

Chapter 2: Materials and Methods.......................................... 9
  Mice
    Plasmid construction

v
Chapter 3: Naïve foreign peptide-specific CD4+ T cell populations are shaped by negative selection on similar self peptides.

Chapter 4: CD4+ T cell persistence and function after infection are maintained
by low-level peptide:MHC class II presentation

Chapter 5: A large population of CD4+ T cells specific for a single p:MHCII plays a large role in control of Salmonella enterica infection

Chapter 6: Conclusions

References
# List of Tables

## Chapter 3

1. Nonamer cores of 20 foreign I-A\(^b\)-binding peptides .............................. 37
2. MHCII anchoring characteristics for p:I-A\(^b\) ................................. 38
### List of Figures

**Chapter 3**

1. Relationship between p:I-A\(^b\)-specific CD4\(^+\) T cell population sizes in unimmunized and immunized mice…………………………………………........ 39
2. Alanine substitution analysis of CD4\(^+\) T cell responses to known foreign epitopes……………………………………………………………………………….. 41
3. Conservation of TCR contact amino acids predicts p:I-A\(^b\) cross-reactivity… 43
4. Cross-reactivity with self p:MHCII affects foreign p:MHCII cell number… 44
5. TCR cross-reactivity with foreign peptides similar to MOG:I-A\(^b\) can elicit EAE…………………………………………………………………………………… 46

**Chapter 4**

6. Intragastric ST-2W infection is controlled by CD4\(^+\) T cells and persists in the MLN……………………………………………………………………………………………… 61
7. ST-2W infection induces a stable population of 2W:I-A\(^b\)-specific T cells…. 63
8. CD4\(^+\) T cell proliferation occurs at the site of persistent infection…….. 65
9. Numerical stability of CD4\(^+\) T cells is dependent on p:MHCII-induced proliferation………………………………………………………………………………. 66
10. ST-2W infection induces multifunctional Th1 cells……………….. 68
11. Function of CD4\(^+\) memory-phenotype T cells is dependent upon antigen abundance………………………………………………………………………….. 70
Chapter 5

12. The response of a single large CD4+ T cell population affects the bacterial load during persistent ST infection.
Chapter 1

Background and Introduction

Immunology began as a scientific discipline with major implications for human health when Edward Jenner demonstrated that inoculation with the infectious material of cowpox could protect against acquisition of the human disease smallpox. He termed this procedure vaccination. Many years later, it was discovered that microbes are the infectious causes of transmissible diseases. Louis Pasteur developed methods to isolate infectious microbes, inactivate them, and inject them. These principles led to effective vaccines against many deadly human diseases, including polio, measles, mumps, rubella, diphtheria, and tetanus. Today, vaccination has saved more lives than any other medical intervention.

It is now appreciated that the vaccine approach originally described by Pasteur successfully produces humoral immunity to extracellular pathogens. This is due to the fact that transient antigen exposure is sufficient to generate long-lived plasma cells that can produce neutralizing antibodies (Slifka, Antia et al. 1998). For intracellular infections, such as *Mycobacterium tuberculosis, Leishmania major,* and *Salmonella enterica,* however, the classical approach to vaccinology has failed because antibodies cannot eliminate microbes that persist within the phagosomes of phagocytes (Tubo and Jenkins 2014). CD4\(^+\) T cells are the critical cell type for immune protection from these types of infections (Erb, Blank et al. 1996, Scanga, Mohan et al. 2000, Mogues, Goodrich
et al. 2001, Reith and Mach 2001). Therefore, a deeper understanding of the factors that create robust and long-lived CD4$^+$ T cell-mediated immunity will be necessary to create new classes of vaccines to target these important human pathogens (Sette and Rappuoli 2010).

**Adaptive immunity and the Clonal Selection Theory**

The clonal selection theory of acquired immunity was proposed by Sir MacFarlane Burnet in the 1950s and was the first comprehensive explanation for how the adaptive immune system responds to foreign antigens. Burnet correctly suggested that individual cells, or clones, each express a unique receptor that determines its specificity. Clones that express receptors with affinity for self antigens are deleted at an early stage of development (clonal deletion). Mature clones that bind cognate antigen through their unique receptor undergo cell activation and division (clonal expansion) and the progeny of a divided cells bear receptors of the same specificity as the parental cells (Burnet 1959). Remarkably, this description of the key principles of the adaptive immune system occurred before lymphocytes were even discovered as the responsible cells (Gowans, McGregor et al. 1962). Thus, while many important details have been added, such as the mechanism for diversification of antigen receptors (Tonegawa, Steinberg et al. 1974, Schatz, Oettinger et al. 1989), clonal selection remains the guiding principle for studying the adaptive immune system.

**Generation and selection of TCR diversity during thymic development**
Burnet’s clonal selection theory predicted that each lymphocyte would have a unique receptor and that receptors with affinity for self antigens would be deleted at an early stage of development. Many details of this process for T cells are now clear. CD4<sup>+</sup> and CD8<sup>+</sup> T cells are generated from hematopoietic progenitors that migrate from the bone marrow to the thymus. Progenitor cells enter the thymus through the bloodstream, proliferate, and pass through a series of developmental stages before exiting back to the blood and secondary lymphoid organs. TCR diversity is established by somatic recombination of gene segments encoded by TCRα and TCRβ gene loci (Nikolich-Zugich, Slifka et al. 2004). This process occurs independently in a random fashion in each developing T cell clone. Variable, Diverse, and Joining (VDJ) gene segments are flanked by 12 and 23 base pair recombination signal sequence (RSS) motifs that are recognized by a complex of enzymes known as the V(D)J recombinase. Lymphocyte-specific expression of the RAG1 and RAG2 enzymes within the recombinase cleaves DNA at RSS sites to create sites for joining the coding ends of V, D, and J segments. Non-templated stretches of nucleotides are added by the enzyme terminal deoxynucleotidyl transferase (TdT), and the gene segments are recombined. This occurs in a stepwise process in each developing T cell clone to produce functional TCRα and TCRβ genes. The multiple combinations of gene pairs and junctional diversities created by this process can produce a theoretical repertoire of 10<sup>15</sup> unique TCRs in mice and 10<sup>18</sup> in humans (Davis 1990). Thus, the process ensures a highly diverse set of clonally distributed T cell receptors (TCRs) that are collectively capable of recognizing a vast number of p:MHC antigens.
Rearranged TCRs are expressed on the cell surface as αβ heterodimers and determine the fate of each clone during subsequent thymic selection. Double positive (DP) thymocytes, characterized by expression of both CD4 and CD8 co-receptors, undergo positive selection if their TCRs are capable of reacting with self peptides and MHC molecules presented by cortical epithelial cells (cTECs) (Jameson, Hogquist et al. 1995). Most clones undergo death by neglect, a form of programmed cell death, however, DP thymocytes that recognize self p:MHCII survive and further differentiate into CD4 single positive (CD4SP) cells (Singer, Adoro et al. 2008). CD4SPs then migrate to the medulla, where they undergo negative selection on ubiquitous and tissue-restricted self p:MHCII expressed by thymic dendritic cells (DCs) and medullary thymic epithelial cells (mTECs) (Anderson, Venanzi et al. 2002, Klein, Kyewski et al. 2014). Thymocytes expressing TCRs with a high affinity for self p:MHCII expressed by DCs and mTECs undergo apoptosis or become FoxP3+ T regulatory cells (Klein, Kyewski et al. 2014). Thus, only T cells with a narrow range of reactivity to self p:MHCII survive positive and negative selection in the thymus.

p:MHCII recognition by CD4+ T cells

Knowledge of p:MHC presentation is necessary for determining the critical aspects of TCR recognition. MHCII molecules are expressed on the surface of antigen presenting cells (APCs) such as macrophages and DCs, which take up extracellular proteins by phagocytosis or macropinocytosis. Internalized proteins are compartmentalized and subjected to proteolytic cleavage by enzymes such as cathepsins.
The resulting peptides that are capable of binding MHCII are transported to the cell surface.

The open-ended structure of MHCII peptide binding grooves permits binding and presentation of peptides from 9-25 amino acids in length (Chicz, Urban et al. 1992, Hunt, Michel et al. 1992). As a result, determining the relevant MHCII binding register, or core nonamer peptide for CD4+ TCR recognition, has been a difficult problem (Wang, Sidney et al. 2008, Chaves, Lee et al. 2012). A solution to this problem has been to combine information gathered from individual p:MHCII crystal structures (Painter and Stern 2012) with peptide elution studies (Falk, Rotzschke et al. 1991, Zhu, Rudensky et al. 2003). This information has been used to generate effective computational methods for predicting peptides that will bind to a multiple MHCII allelomorphs based upon the presence of key MHCII anchor residues within the peptide (Chaves, Lee et al. 2012).

TCRs bind p:MHC through 6 complementary determining region (CDR) loops: CDR1α, CDR1β, CDR2α, CDR2β, CDR3α, and CDR3β. The CDR1 and CDR2 loops consist of germline encoded V regions that bind to conserved regions of MHC molecules (Marrack, Scott-Browne et al. 2008). These interactions explain the inherent MHC reactivity of highly diverse TCRs. They can also account for the common diagonal docking angle of TCRs with respect to the axis of the MHC peptide binding groove (Rudolph, Stanfield et al. 2006). This orients the CDR3, or hypervariable, loops of the alpha and beta chains over the MHC-bound peptide to confer TCR specificity. The affinities of TCRs for p:MHC is typically in the 1-100 µM range (Davis 1998). Thus, this interaction is thought to be relatively low affinity but highly specific. More recent
The pre-immune T cell repertoire

The pre-immune repertoire, or naïve repertoire, consists of millions cells that have exited the thymus but never encountered a foreign p:MHCII agonistic ligand. CD4⁺ T cells predominantly exist in a naïve or quiescent state until encountering cognate p:MHCII through their TCR. Naïve T cells in mice are marked by low expression of the cell surface marker CD44 and express the trafficking molecules L-selectin (CD62L) and CCR7, which allow them to enter lymphatic tissues via high endothelial venules and migrate along a CCL19 and CCL21 chemokine gradient.

The mouse repertoire has an estimated 2x10⁶ different αβ TCR clonotypes and the human repertoire contains approximately 10 times more (Arstila, Casrouge et al. 1999, Casrouge, Beaudoin et al. 2000). A completely naïve repertoire would only persist under completely antigen-free conditions. Under specific pathogen free (SPF) conditions, the mouse repertoire contains approximately 70% naïve phenotype cells, 10% recent thymic immigrants, 10% memory phenotype cells, and 10% FoxP3⁺ natural T regulatory cells (Jenkins, Chu et al. 2010). However, it has recently been shown that the ratio of circulating memory cells to naïve cells is higher in humans than the commonly studied SPF mice. This is true for even p:MHCII-specific populations of cells that recognize foreign peptides to which the host has not been exposed, suggesting TCR cross-reactivity.
for environmental microbes could alter the “pre-immune” repertoire in ways not yet known (Su, Kidd et al. 2013).

**Immunodominance**

Although a pathogen can theoretically express thousands of peptides during infection, only a fraction of peptides induce robust immune responses. This concept is referred to as immunodominance. Predicting immunodominant T cell responses has been the focus of considerable efforts and it is now clear that numerous factors contribute to this phenomenon. The amount and duration of p:MHCII antigen presentation are CD4+ T cell extrinsic signals that affect clonal expansion and differentiation (Itano, McSorley et al. 2003, Obst, van Santen et al. 2005, Prlic, Hernandez-Hoyos et al. 2006, Henrickson, Mempel et al. 2008, Tubo, Pagan et al. 2013). Foreign protein expression levels and location, APC proteolysis, and MHCII binding stability dictate the amount of p:MHCII produced, and thus the strength of these signals (Sant, Chaves et al. 2005, Yewdell 2006). The composition of the naïve T cell repertoire also shapes immunodominance. Evidence for this comes from studies utilizing p:MHC tetramer-based cell enrichment methods to quantify rare, polyclonal p:MHC-specific CD4+ and CD8+ T cells in both unexposed and antigen-experienced subjects. This work has demonstrated that the number of naïve cells in the naïve repertoire is perhaps the most critical predictor of immunodominant T cell responses (Moon, Chu et al. 2007, Kotturi, Scott et al. 2008, Obar, Khanna et al. 2008, Flesch, Woo et al. 2010, Schmidt, Neumann-Haefelin et al. 2011, Tan, La Gruta et al. 2011, Kwok, Tan et al. 2012).
**Statement of thesis**

This thesis focuses on how foreign p:MHCII-specific CD4\(^+\) T cell populations are generated as naïve cells and maintained as memory cells following infection. A p:MHCII tetramer-based cell enrichment strategy was employed to demonstrate that TCR cross-reactivity for different MHCII-bound peptides is largely dependent upon amino acids at TCR contact positions, that the sizes of naïve CD4\(^+\) T cell populations recognizing foreign pMHCII ligands is influenced by TCR cross-reactivity with MHCII-binding self peptides, and that prolonged foreign p:MHCII presentation maintains CD4\(^+\) T cell stability and function following infection. This work is significant because there are no current vaccines that successfully target and sustain long lasting CD4\(^+\) T cell-mediated immunity.
Chapter 2

Materials and Methods

Mice

C57BL/6 (B6), B6.SJL-Ptprc<sup>a</sup>Pep3<sup>b</sup>/BoyJ (CD45.1) mice, and 129SvJ (129) female mice were purchased from the Jackson labs. Act-2W, (B6 x 129) F<sub>1</sub>, (CD45.1 x 129) F<sub>1</sub>, and (Act-2W x 129) F<sub>1</sub> mice were bred in our facilities. All mice were housed in specific pathogen-free conditions at the University of Minnesota and all experiments were conducted in accordance with institutional and federal guidelines.

Plasmid construction

Sequences encoding pRMHa-3 vectors containing the α- and β-chains of I-A<sup>b</sup> under the control of the metallothionein promoter were constructed as described in Moon et al. (Moon, Chu et al. 2007). Sequences encoding antigenic peptides (ESAT6: QQWNFAGIEAAASA (Reiley, Shafiani et al. 2010); FliC: VQNRFSAITNLGNT (Moon, Chu et al. 2007); STM1540-3: VYYTTYAPQAT; STM1540-1: YTTYAPQATSA; eGFP: HDFFKSAMPEGYVQE; OVA3C: GHAAHAEEINAG (Robertson, Jensen et al. 2000, Moon, Dash et al. 2011); RpIF: VFVSPAHHID (Li and McSorley 2013); OVA2C: QAVHAHAEIN (Robertson, Jensen et al. 2000, Moon, Dash et al. 2011); 3K: EAQKAKANKAVDKA; Calnex: LVVKPNAAHHAIS; Aasf: GVSSPAVQESI (Li and McSorley 2013); 1G1W: EAGGALANWAVDSA (Chu, Moon et al. 2010); LLO: NEKYAQAYPNVS (Geginat, Schenk et al. 2001); Derp1:
CQIYPNPNVKI; MP98: HQYMTALSNEVF; PmpG-1: YVDPAAGG (Li and McSorley 2013); CTB: NNKTPHAIAAIS (Cong, Bowdon et al. 1996); GP66: DIYKGVYQFKSV (Oxenius, Bachmann et al. 1995, Moon, Dash et al. 2011); 2W: EAWGALANWAVDS; CD4Ag28m: VEIHRPVPGA (Grover, Blanchard et al. 2012); 2W109: EYWGPLPNWVVD; and MOG: GWYRPSFVRV were fused to the N terminus of the I-Aβ β-chain via a flexible polyglycine linker. For LLO:I-Aβ, OVA2C:I-Aβ, OVA3C:I-Aβ, STM1540-3:I-Aβ, STM1540-1:I-Aβ, 2W109:I-Aβ, and MOG:I-Aβ tetramers, a disulfide trap strategy was used to prevent peptide movement within the MHCII groove and lock in the desired peptide binding register, as previously described (Stadinski, Zhang et al. 2010, Moon, Dash et al. 2011).

**p:MHCII tetramer production**

Tetramers were designed produced in the Jenkins lab and by collaborators (Li and McSorley 2013). Peptide:I-Aβ molecules were expressed in *Drosophila melanogaster* S2 cells using the Drosophila Expression System kit (Invitrogen, Carlsbad, CA). The soluble p:I-Aβ heterodimers were then purified from the cell culture supernatant via nickel affinity chromatography as previously described (Moon, Chu et al. 2007), followed by an additional purification on a Pierce Monomeric Avidin UltraLink Resin (Thermo Scientific, Rockford, IL, USA). The biotinylated p:I-Aβ molecules were eluted with free biotin and washed using Amicon Ultra-15 30kD concentrating filters (Merck Millipore Ltd., Tullagreen, Carrigtwohill Co. Cork, IRL) to remove excess free biotin. The
resulting product was then used to generate tetramers with streptavidin (SA)-phycoerythrin (PE) or (SA)-allophycocyanin (APC) (Prozyme, San Leandro, CA, USA).

**Cell enrichment and flow cytometry**

Chapter 3: Spleen and lymph node cells were harvested and made into single cell suspensions. Tetramer staining was for 1 hour at room temperature with p:IA\textsuperscript{b}-streptavidin-PE and p:IA\textsuperscript{b}-streptavidin-APC tetramers. Anti-PE and anti-APC magnetic beads were then incubated on ice and the samples were enriched for bead-bound cells as previously described (Moon, Chu et al. 2007, Tubo, Pagan et al. 2013). Cells contained within the bound fraction were stained on ice with antibodies specific for the cell surface markers B220 (RA3-6B2), CD11b (MI-70), CD11c (N418), Thy1.2 (53-2.1), CD8\textsuperscript{a} (5H10; Invitrogen, Carlsbad, CA, USA), CD4 (RM4-5), CD3\textepsilon (145-2C11), and CD44 (IM7), each conjugated to a different fluorochrome. All antibodies were from eBioscience (San Diego) unless noted. Cells were then analyzed on an LSR II or Fortessa (Becton Dickinson) flow cytometer. Data were analyzed with FlowJo (TreeStar, Ashland, OR, USA).

Chapter 4: Spleen and lymph node cells were harvested and made into single cell suspensions. 10% of the cell suspensions was removed for plating bacterial CFUs with the addition of 0.1% Triton. The remaining 90% of the samples were stained for 1 hour at room temperature with or 2W:IA\textsuperscript{b}-streptavidin-APC tetramer, enriched for tetramer bound cells, counted and labeled with antibodies as previously described (Moon, Chu et
al. 2007). Cells events were collected on an LSRII or Fortessa flow cytometer (Becton Dickinson) and analyzed using FlowJo software (TreeStar).

**Peptide immunization**

Peptides were dissolved in PBS and emulsified in Complete Freund’s Adjuvant (CFA). Fifty microliters of emulsion containing a final concentration of 10 µg peptide was injected subcutaneously at the base of the tail.

**PSSM**

A position-specific scoring matrix (PSSM) for I-A\(^b\) binding was generated by using the I-A\(^b\) binding motif described by Zhu et al. (Zhu, Rudensky et al. 2003) to align an additional 128 I-A\(^b\)-bound peptides (Dongre, Kovats et al. 2001, Karunakaran, Rey-Ladino et al. 2008).

**ELISPOT**

Mice were immunized by injection of 10 µg peptide emulsified in CFA, as described above. Two weeks later, CD4\(^+\) T cells were isolated from the draining inguinal and para-aortic lymph nodes and purified using positive selection with anti-CD4 magnetic beads (Miltenyi). The cells were challenged with candidate peptides *in vitro* at a final concentration of 20 µM and analyzed for the presence of IFN\(\gamma\)-producing CD4\(^+\) T cells by ELISPOT. The ELISPOT assay was adapted from a published protocol (Streeck, Frahm et al. 2009) and assay guidelines from a commercially available kit (Mabtech).
Individual wells (peptides) were performed in triplicate. Each sample of cells included negative controls consisting of un-stimulated cells as well as cells stimulated with an irrelevant peptide. Positive controls included cells stimulated with ConA.

**Alanine scanning approach**

Alanine substitutions were made at P-1 to P10 for 14 foreign peptides. In places where A was in the native sequence, V was substituted, with the exception of P4 or P6, where P was substituted. This was to avoid shifting the I-A\(^b\)-binding register because P is the preferred amino acid at P4 and P6 based on a peptide-I-A\(^b\)-binding matrix. The peptides were tested as stimulators of a secondary *in vitro* response of bulk lymph node cells primed with peptide emulsified in CFA. The results were analyzed using the publicly available script “heatmap.2” that is available for the statistical analysis program “R.”

**Determination of number of overlapping peptides in known bacterial proteomes**

The set of 1,632 complete bacterial proteomes was downloaded in fasta format from Uniprot (www.uniprot.org) using the publicly available batch retrieval perl script. Each of these proteomes was then parsed into all possible ten amino acid substrings and I-A\(^b\)-binding scores for each substring were calculated. The predicted I-A\(^b\)-binding peptides from each bacterial proteome were taken as those with I-A\(^b\)-binding scores greater than the 1st percentile cutoff for mouse peptides, as determined above. Peptides of interest were then screened against the bacterial peptides above the 1% I-A\(^b\)-binding threshold for overlap at amino acid positions 2, 5, 7, and 8 using custom c++ scripts.
Determination of number of overlapping peptides in mouse proteome

The *Mus musculus* proteome was downloaded from the NCBI website (ftp://ftp.ncbi.nlm.nih.gov/genomes/M_musculus/protein/protein.fa.gz). All possible ten amino acid substrings were extracted from these peptide sequences and each substring was given a score based on its homology with known I-A\(^b\)-binding peptides as published previously (Zhu, Rudensky et al. 2003). Using the distribution of calculated I-A\(^b\)-binding scores, the 1\(^{st}\) and 5\(^{th}\) percentile predicted I-A\(^b\)-binding score cutoff for ten amino acid peptides was determined. We then utilized custom c++ scripts to compare ten amino acid foreign peptides of interest against the top 5% I-A\(^b\)-binding mouse peptides and count the number of peptides with identical amino acids at positions 2, 5, 7, and 8.

EAE Induction and clinical evaluation

For determining whether the bacterial expressed peptides similar to MOG could induced active EAE in B6 mice (Figure 5F), mice were immunized subcutaneously with 100 µg of MOG39–49 or 11 amino acid versions of homologous foreign peptides emulsified in CFA containing 4 mg/ml *Mycobacterium tuberculosis*. In addition, the mice received 200 ng of pertussis toxin (List Biological Laboratories) intravenously on the day of, and 2 days after, peptide immunization. Individual animals were graded according to clinical severity as follows: grade 0, no abnormality; grade 1, limp tail; grade 2, limp tail and hind limb weakness; grade 3, partial hind limb paralysis; grade 4, complete hind limb paralysis; as previously described (Fife, Huffnagle et al. 2000).
Salmonella infections

Mice were pre-treated by 12-hour food deprivation or intragastric gavage with 100 µl 5% sodium bicarbonate solution, pH 9.0 prior to infection. They were then given $10^8$ S. enterica serovar Typhimurium (ST) strain SL1344 or recombinant S. enterica serovar Typhimurium strain SL1344 OmpC-2W (ST-2W, produced as described below) by intragastric gavage. Enrofloxacin was included in the drinking water at 2 mg/ml in some cases.

Production of the ST-2W strain

ST was tagged chromosomally with the 2W peptide (EAWGALANWAVDSA) as previously described by Uzzau et al. 2001. Briefly, primers were designed with extension arms homologous to the 3’ portion of the OmpC gene, deleting the stop codon, and extending downstream from it. Additionally, a single FLAG sequence was included, for blotting purposes, before reintroduction of the stop codon while kanamycin resistance was introduced downstream of the newly incorporated stop codon. PCR products were generated by amplification from a template plasmid (pJM1) encoding the 2W peptide and FLAG epitope. PCR products were used directly for electrottransformation into ST containing the temperature-sensitive pKD46 plasmid, carrying arabinose inducible bacteriophage λ red genes. Bacterial suspensions in 10% glycerol were mixed with 0.5–1 μg of PCR product and incubated on ice for 30 min before transferring to a chilled 0.2-cm cuvette. Cuvettes were subjected to a single pulse of 12.5 kV/cm. After recovering for 1
hour at 37°C in SOC medium, bacteria were plated on LB agar plates supplemented with 50 µg/ml kanamycin. DNA sequencing was used to verify recombination.

**Western blotting**

Wild type and ST-2W bacteria were grown overnight in LB Broth. Bacteria were centrifuged and pellets were suspended in 1 ml of 1X CelLytic™ B cell lysis reagent (Sigma), containing 0.2 mg/ml lysozyme (RPI Corp.), 50 U/ml Benzonase® Nuclease (Sigma), and 25 µl of Protease Inhibitor Cocktail (Sigma). Cell extracts were centrifuged and supernatants containing soluble proteins were removed for SDS-PAGE analysis. Separated proteins were transferred to a nitrocellulose membrane then blocked in 5% BSA for 1 hour. Membranes were probed with mouse anti-FLAG M2 monoclonal antibody (Sigma) at 1 µg/ml in 1% BSA for 1 hour at room temperature with gentle agitation. The blot was washed and then probed with 1 µg/ml of a secondary goat anti-mouse conjugated to AlexaFluor-680 (Invitrogen) for 1 hour at room temperature. The blot was analyzed on an Odyssey® Imaging System (Li-Cor) at 700 nm.

**BrdU incorporation studies**

BrdU (Sigma-Aldrich) was dissolved in PBS at a concentration of 10 mg/ml. In vivo BrdU labeling was performed by intraperitoneal injection of 1 mg of BrdU on day 50 of ST-2W infection. 12 hours later, spleens and MLN were harvested individually and BrdU incorporation into DNA was detected by intracellular staining according to the manufacturer’s specifications (BD Biosciences).
**Lymphokine production**

ST-2W-infected mice were injected with 100 µg of 2W peptide. Two hours later, spleens and mesenteric lymph node cells were harvested in media containing 10 µg/ml brefeldin A. The resulting cell suspensions were fixed and permeabilized and stained with IFN-γ, TNF, and IL-2 antibodies as previously described (Pepper, Linehan et al. 2010).

**Cell transfer experiments**

CD4⁺ T cells were isolated from spleens and lymph nodes of (CD45.1 x 129) F₁ mice infected 50 days earlier with ST-2W and purified using Miltenyi CD4 isolation kits. The cells were then washed in EHAA media, suspended to a concentration of 5 x 10⁷ cells/ml, incubated with 5 µM CFSE (Invitrogen) at 37°C for 10 minutes, and washed with pre-warmed EHAA media before intravenous injection into (B6 x 129) F₁ hosts. For cytokine experiments, purified CD4⁺ T cells were washed and injected intravenously without CFSE labeling. To control for experiment-to-experiment variability, the numbers of parked cells were normalized for the input number of donor-derived CD4⁺ T cells.

**Statistical analysis**

Statistical differences between normally distributed data sets were assessed in most cases using paired or unpaired two-tailed Student’s t tests with Prism (Graphpad) software. A Mann-Whitney test was used in one case where the data were not normally distributed (Fig. 9E).
Naïve foreign peptide-specific CD4$^+$ T cell populations are shaped by negative selection on similar self peptides

3.1 Introduction

Naïve CD4$^+$ and CD8$^+$ T cell populations vary in size over several orders of magnitude (Moon, Chu et al. 2007, Kotturi, Scott et al. 2008, Obar, Khanna et al. 2008, Flesch, Woo et al. 2010, Schmidt, Neumann-Haefelin et al. 2011, Tan, La Gruta et al. 2011, Kwok, Tan et al. 2012). It remains unclear, however, which factors account for this variability. The chemistry of peptide:major histocompatibility complex (p:MHC) combinations, the ligands for T cell antigen receptors (TCRs) could be one important determinant (Turner, Kedzierska et al. 2005, Chu, Moon et al. 2010). Peptides containing large hydrophobic amino acids that protrude from the MHC-binding groove appear to have a greater propensity to form stable TCR contacts than peptides with less featured amino acids. Another possibility is that foreign p:MHC-specific T cell populations with a high degree of cross-reactivity for self p:MHC encountered during thymic development may be pared by negative selection. This possibility, however, has been difficult to study because the rules for TCR cross-reactivity have not been worked out.

Recent progress, however, has been made in this area for p:MHCII-specific TCRs expressed by CD4$^+$ T cells (De la Herran-Arita, Kornum et al. 2013, Su, Kidd et al. 2013, Birnbaum, Mendoza et al. 2014). Peptides of up to 20 amino acids bind to MHCII
molecules via a linear nine amino acid core sequence (Rudolph, Stanfield et al. 2006). Certain amino acids at positions within that stretch, often at positions 1, 4, 6 and 9, anchor the peptide by fitting into pockets in the MHCII groove (Painter and Stern 2012). Because each pocket prefers one of several amino acids, certain MHCII molecules have the capacity to bind as many as 10% of all peptides. The amino acids at positions 2, 3, 5, and 8 generally point up out of the MHCII binding groove. Structural biology studies indicate that complementarity determining region (CDR) 1 and 2 of the TCR interact with the MHCII alpha helices, while CDR3 interacts mainly with the up pointing residues in the peptide (Marrack, Scott-Browne et al. 2008). Using soluble TCRs to probe large yeast-displayed MHCII-bound peptide libraries, Birnbaum, et al., found that a single TCR could bind dozens of peptides with similar TCR contact amino acids (Birnbaum, Mendoza et al. 2014) but very different MHCII anchor amino acids. If TCR specificity is determined by only 4 of the 9 amino acids in an MHCII-bound peptide, 1 in 160,000 \((20^4)\) peptides that bind the same MHCII molecule will have the same TCR contact amino acids and thus bind to the same TCR. Because a mammalian proteome can in theory produce \(1.5\times10^6\) different MHCII-binding peptides, any given MHCII-binding foreign peptide should have on average about 10 MHCII-binding self peptide homologs \((1.5\times10^6/1.6\times10^5)\). These homologs could reduce the size of the foreign peptide-specific population by causing negative selection. Since different foreign peptides could have different numbers of self-peptide homologs, the corresponding naïve cell populations could undergo different degrees of selection and vary in size for this reason.
We tested this premise in C57BL/6 (B6) mice, which express the I-A<sup>b</sup> MHCII molecule. We confirmed that TCR specificity for peptides bound to I-A<sup>b</sup>(p:I-A<sup>b</sup>) depended primarily on four TCR contact amino acids and was relatively independent of the nature of the MHCII anchor residues. This type of TCR cross-reactivity reduced the size of T cell populations specific for foreign peptides with many self peptide homologs in the host proteome. In addition, a self peptide from a tissue restricted protein had a large naïve population that showed evidence of minimal evidence of negative selection and could be primed with bacterial peptides that shared only 4 TCR contact amino acids. Thus, TCR cross-reactivity between similar self and foreign peptides influenced the composition of the foreign peptide-specific T cell repertoire via negative selection, and accounted for the capacity of microbial peptides to trigger autoimmunity to an ignored self peptide.

3.2 Results

**Foreign-p:I-A<sup>b</sup>-specific naïve T cell populations vary in size.**

Earlier work indicated that naïve CD4<sup>+</sup> T cell populations specific for 3 different I-A<sup>b</sup>-bound foreign (non-mouse) peptides varied in size in B6 mice (Moon et al. 2007). We enumerated many more populations to get a better sense of the normal range and to set the stage for assessment of the effect of negative selection by related self (mouse) peptides. A p:MHCII tetramer-based cell enrichment and flow cytometry strategy (Moon, Chu et al. 2009) was used to identify naïve CD4<sup>+</sup> T cells specific for 20 I-A<sup>b</sup>-bound
foreign peptides (Table 1) in B6 mice that were never exposed to these peptides (Moon, Chu et al. 2007). In each case, cells from the secondary lymphoid organs were stained with a pair of I-A<sup>b</sup> tetramers containing the same peptide but labeled with different fluorochromes, and then magnetic beads. Labeled cells were enriched on magnetized columns and stained with antibodies specific for informative surface markers. The foreign p:I-A<sup>b</sup>-specific cells in the bound fraction were detected by gating on lymphocyte-sized, single cells that were stained with CD90.2 but not B220, CD11b, or CD11c, and with CD4 but not CD8 (Figure 1A). The double tetramer staining strategy was used for each of the populations to enhance the sensitivity and TCR-specificity of the assay (Stetson, Mohrs et al. 2002, Tubo, Pagan et al. 2013). Representative staining patterns for three different I-A<sup>b</sup>-binding foreign peptides are shown in Figure 1B (left columns). Most double tetramer-binding cells in each population had the CD4<sup>+</sup>CD90.2<sup>-</sup> phenotype of naïve cells, indicating that the mice had no prior exposure to these epitopes. As quantified in Figure 1C (filled circles), the absolute number of naïve CD4<sup>+</sup> cells ranged over two orders of magnitude, from fewer than an average of 5 ESAT6<sub>8-16</sub>:I-A<sup>b+</sup> cells per mouse to over 500 CD4Ag28m<sub>608-616</sub>:I-A<sup>b+</sup> cells per mouse.

Mice were then immunized with 10 µg of the 20 peptides emulsified in complete Freund’s adjuvant (CFA) to test the hypothesis that naïve cell number predicts the magnitude of the effector cell response. Immunization caused clonal expansion and CD44 induction on the relevant p:I-A<sup>b</sup> specific T cell populations for all 20 peptides two weeks after immunization (Figure 1B, right columns). The hierarchy of effector cell generation corresponded to the number of naïve cells in most cases (Figure 1C). Across the 20 CD4<sup>+</sup>...
T cell populations queried, naïve cell number predicted 65% of the variance in the
number of effector cells generated in response to peptide immunization (Figure 1D).
These results suggest that the naïve T cell numbers shown in Figure 1C were accurate
and solidified the conclusion that the naïve cell number predicts the magnitude of the
effector cell response.

**Determining the nonamer core peptides for p:I-A^b**

It was next necessary to identify the TCR contact amino acids for the 20 foreign
peptides to make it possible to search for mouse peptides with matching residues. The
strategy was to identify the amino acids in the foreign peptides that bound to I-A^b, which
by default would identify the other amino acids as up pointing TCR contacts. Zhu *et al.*
(Zhu, Rudensky *et al.* 2003) had already identified an I-A^b binding motif for nonamer
peptides by aligning the sequences of 78 peptides eluted from I-A^b. We used this motif to
align an additional 128 I-A^b-bound peptides (Dongre, Kovats *et al.* 2001, Karunakaran,
Rey-Ladino *et al.* 2008) (Table 2A) with the goal making the motif more accurate. We
assumed that a correct alignment would reveal certain preferred amino acids at positions
in the peptide that occupied I-A^b binding pockets. Indeed, the aligned peptides contained
tyrosine, tryptophan, for phenylalanine at position (P) 1, proline at P4, proline or alanine
at P6, and asparagine at P7 at frequencies that were at least 3 times greater than predicted
by the average overall abundance of these amino acids in organisms. This result is
consistent with the interpretation that the side chains of the P1, 4, 6, and 7 amino acids
occupy pockets in I-A^b as proposed by other investigators (Liu, Dai *et al.* 2002, Zhu,
Rudensky et al. 2003). In contrast, P2, 3, 5, 8, and 9 of the aligned peptides did not show greater than 3-fold enrichment for any amino acid suggesting that these residues do not play major roles in I-A\textsuperscript{b} binding, although structural biology studies indicate that P9 plays a minor role (Zhu, Rudensky et al. 2003). These results suggest that peptides bind to I-A\textsuperscript{b} mainly by certain amino acids at P1, 4, 6, and 7.

The data in Table 2B was then used to identify the most likely I-A\textsuperscript{b}-binding register for each of the 20 foreign peptides in the I-A\textsuperscript{b} tetramers used in Figure 1. For peptides that were longer than 9 amino acids, the possible nonamers were given a score based on the sum of the frequency values for the 9 amino acids as determined from Table 2B. In each case, one of the possible nonamers had a much higher score than the others and was chosen as the I-A\textsuperscript{b}-binding core for that peptide (Table 1).

If the I-A\textsuperscript{b} binding registers shown in Table 1 are correct, then P5 should be the most important amino acid for TCR recognition because the CDR3 alpha and beta regions of TCRs generally interact with this amino acid (Marrack, Scott-Browne et al. 2008). An alanine substitution scan was performed to test this possibility. All alanine substituted peptides were predicted to bind I-A\textsuperscript{b} binding because alanine occurred often at P1, 6, 4, and 7 in the set of 206 I-A\textsuperscript{b}-bound peptides shown in Table 2. Because it was not clear whether amino acids flanking the I-A\textsuperscript{b}-bound nonamers could serve as additional TCR contacts, mice were primed with 11 amino acid versions of 13 of the 20 foreign peptides. Two weeks later, CD4\textsuperscript{+} T cells from the draining lymph nodes were stimulated with the priming peptide, or peptides with single alanine substitutions within the nonamer core plus as well as the P-1 and P10 N- and C-terminal flanking residues. The T cell
response was then measured by the interferon-γ Enzyme-Linked ImmunoSpot (ELISPOT) assay. Alanine substitutions at the predicted P5 residues reduced reactivity by parent peptide-primed T cells by an average of 90% (Figure 2A). Substitutions at P2, 3, or 8 also reduced T cell reactivity by at least 60% as predicted for TCR contact residues (Huseby, White et al. 2005). In contrast, substitutions at P1 or 9, as well as the flanking amino acids, reduced T cell reactivity on average less than 30%. These results are consistent with P2, 3, 5, and 8, but not P1, P9, or flanking amino acids being TCR contact residues for I-A^b-bound peptides.

It was surprising, however, to find that substitutions at P4, P6, and P7 reduced cell reactivity by parent peptide-primed T cells because Table 2B and structural biology experiments (Liu, Dai et al. 2002, Zhu, Rudensky et al. 2003) indicate that these positions are I-A^b anchors. Because there was some variability in the effect of alanine substitution across the 13 peptides as evidenced by the large standard deviations in Figure 2A, an unsupervised clustering of amino acid positions was performed as an independent way to identify the important TCR contact amino acids (Figure 2B). As expected, P2, 5, and 8 clustered for large effects on TCR recognition. Notably, P7 clustered with these amino acids better than P3. P1, 4, 6, and 9 clustered for smaller effects on TCR recognition. These results and those in Table 2B suggest that peptides are anchored to I-A^b by primarily by amino acids at P1, 4, 6, and 7, with a small contribution from P9 and that amino acids at P2, 5, 7, and 8 are the main TCR contacts with a smaller contribution from P3. The behavior of P7 as an I-A^b anchor and TCR contact is consistent with the fact that
it occupies a shallow pocket in I-A\(^b\) allowing amino acid side chains to protrude for recognition by TCRs (Zhu, Rudensky et al. 2003).

**TCR contact amino acid conservation predicts p:MHCII cross-reactivity within polyclonal CD4\(^+\) T cell populations**

The knowledge that amino acids at P2, 3, 5, 7, and 8 are the TCR contacts for I-A\(^b\)-bound peptides made it possible to test the limits of cross-reactivity for I-A\(^b\)-restricted TCRs. Over a 1,000 bacterial proteomes were searched for predicted I-A\(^b\)-binding nonamer peptides that matched the 3K, LLO, GP66, and 2W peptides at 4 of the P2, 3, 5, 7, or 8 amino acids but at most one of the P1, 4, or 9 amino acids. Three such peptides were identified for 3K, 1 for LLO, and 5 for GP66 and 2W. B6 mice were then primed with the parent 3K, LLO, GP66, or 2W peptides, or the corresponding variant peptides and the number of responding CD4\(^+\) T cells was measured with 3K, LLO, GP66, or 2W:I-A\(^b\) tetramers. If TCR cross-reactivity existed between a parent peptide and a variant peptide, then priming with the variant peptide should induce the clonal expansion of at least a subset of the T cells specific for the parent p:I-A\(^b\) ligand. As expected, mice primed 11 days earlier with 3K, LLO, GP66, or 2W peptides in CFA had large populations of CD44\(^{high}\) cells that bound to the parent p:I-A\(^b\) tetramer (Figure 3A). Priming with the parent peptide caused this clonal expansion as evidenced by the fact that mice primed with CFA alone had about the same number of parent p:I-A\(^b\)-specific cells as naïve mice. The cells in mice primed with CFA alone were also phenotypically naïve based on low expression of CD44, with the exception of the 2W:I-A\(^b\)-specific population,
which had a small population of \(CD_{44}^{\text{high}}\) cells. Remarkably, 1 of 3 3K, 1 of 1 LLO, 2 of 5 GP66, and 4 of 5 2W variant peptides primed significant expansion of \(CD_{44}^{\text{high}}\) CD\(4^+\) T cells capable of binding the relevant parent p:I-A\(^b\) tetramer (Figure 3B). Three of the 2W variant peptides stimulated the expansion of 2W:I-A\(^b\)-specific cells despite sharing no I-A\(^b\) anchor residues with the 2W peptide. Thus, 8 of 14 (57%) of I-A\(^b\)-binding peptides sharing 4 TCR contact amino acids but at most one I-A\(^b\) anchor residue exhibited significant TCR cross-reactivity.

**Self-p:MHCII cross-reactivity limits naïve population size and responsiveness to homologous foreign-p:MHCII**

The findings in Figure 3 raised the possibility that TCR cross-reactivity between mouse and non-mouse peptides could reduce the size of non-mouse peptide-specific naïve populations in mice due to negative selection. A proof of principle experiment was performed with the 2W peptide, a variant called to 2W109 with the same amino acids at P2, 3, 5, 7, and 8 as the 2W peptide but different residues at P1, 4, 6, and 9 (Figure 4A), and a transgenic mouse expressing the 2W peptide as a self antigen. The 2W epitope is a foreign peptide in B6 mice, and is recognized by a diverse population of approximately 300 naïve CD\(4^+\) T cells (Figure 1A) (Moon, Chu et al. 2007). We have previously shown that expression of 2W peptide as a self-peptide under the control of the actin promoter in *Act-2W* transgenic mice induces tolerance and leads to deletion of 2W:I-A\(^b\)-reactive cells (Moon, Dash et al. 2011).
Naïve 2W109:I-A^b-specific T cells were enumerated with 2W109:I-A^b tetramer in B6 and Act-2W mice to assess whether TCR cross-reactivity could cause negative selection. As shown in Figure 4B, B6 mice contained a large population of approximately 1,000 2W109:I-A^b-specific naïve CD4^+ T cells in the spleen and lymph nodes. In contrast, Act-2W mice had only 400 2W109:I-A^b-specific naïve cells. The reduction was likely due to deletion of 2W:I-A^b/2W109:I-A^b cross-reactive clones, as there was no difference in the number of LLO:I-A^b-specific naïve cells in the two strains.

An ELISPOT approach was also used to study the fine specificity of the 2W109:I-A^b-specific cells that were not deleted in mice expressing 2W:I-A^b. The hypothesis was that the population of 2W109:I-A^b-specific cells in B6 mice contains two subpopulations; one with TCRs that bind the TCR contact amino acids in a geometry that is not affected by the I-A^b anchor residues, and one with TCRs that bind the TCR contact amino acids in a geometry that is. The cross-reactive subpopulation would be deleted in Act-2W mice, while the non-cross-reactive one would remain. The prediction would be that the remaining non-cross-reactive 2W109:I-A^b-specific population would be sensitive to substitutions at I-A^b anchor positions since it is focused on an I-A^b anchor position-dependent conformation of the TCR contacts. In B6 mice, the reactivity of 2W109:I-A^b-specific CD4^+ T cells was reduced by at least 50% when stimulated with peptides containing substitutions at P2, 5, or 8 (Figure 4C, black bars). In contrast, in Act-2W mice, the reactivity of the 2W109:I-A^b-specific population was reduced by at least 50% when stimulated with peptides containing substitutions at P2, 3, 4, 5, 6, 7, or 8 (Figure 4C, white bars). Thus, the 2W109:I-A^b-specific population that survived negative
selection in the presence of 2W:I-A$^b$ appeared to be specific for TCR contact amino acids in an I-A$^b$ anchor residue-dependent fashion.

These results demonstrate that a foreign peptide-specific population can be reduced in size because of deletion of cells that cross-react on a self peptide with the same TCR contacts but different MHCII anchors. Furthermore, the surviving cell population had a signature of exposure to the self peptide of remarkable dependence on the character of the MHCII anchor residues. These results raised the possibility that small foreign peptide-specific populations are small because of deletion of cells with TCRs that are cross-reactive on self peptides with the same TCR contacts but different MHCII anchors. If so, then the expectation would be that the specificity of these populations would be more dependent on the character of the MHCII anchor residues than large populations. Indeed, a significant inverse correlation was observed between the number of naïve cells in a foreign peptide-specific population and the average effect of substitutions at I-A$^b$ anchors P1, 4, 6, and 7 on the reactivity of T cells in that population (Figure 4D).

To further test the hypothesis that naïve populations specific for foreign peptides are shaped by negative selection on self-peptides with similar TCR contact amino acids, we searched for self-peptides that could have contributed to negative selection for each of the foreign p:I-A$^b$-specific CD4$^+$ T cell populations. A bioinformatic screen was developed to identify mouse peptides that were predicted to bind I-A$^b$ and share TCR contact amino acids with the non-mouse peptides of interest. To be inclusive for any self p:I-A$^b$ that could have an effect, the stringency of the search was reduced to include any
combination of 4/5 identical amino acid residues at P2, 3, 5, 7, and 8. Mouse peptide homologs were identified for all the foreign peptides, with a range of 3 to 181. The complete list of mouse peptide homologs along with the associated gene names can be found in Supplemental Figure 1. Importantly, a significant inverse correlation was observed between naïve CD4⁺ T cell population size and the number of self peptide homologs (Figure 4E), although the r² value was only 0.27. These results support the conclusion that TCR cross-reactivity on self peptides plays some role in determining the size of MHCII-bound foreign peptide specific CD4⁺ T cell populations.

A tissue restricted self p:MHCII is recognized by a large naïve T cell population with minimal evidence of negative selection

The results indicated that negative selection on a self peptide with the same TCR contact amino acids could reduce the size of a foreign peptide-specific CD4⁺ T cell population. We considered the corollary case in which a self peptide-specific CD4⁺ T cell population that had not undergone complete deletion could be triggered to cause autoimmunity by a foreign peptide with the same TCR contact amino acids. The self (mouse) peptide of interest was peptide 31-50 from myelin oligodendrocyte glycoprotein (MOG), a brain protein which causes experimental acute encephalomyelitis (EAE) in B6 mice when injected with CFA and pertussis toxin (Miller, Karpus et al. 2010). Limited presentation of I-Aᵇ-bound MOG peptide in the thymus probably explains why not all I-Aᵇ-bound MOG peptide-specific T cells are deleted (Ben-Nun, Kerlero de Rosbo et al. 2006). We produced an I-Aᵇ tetramer containing MOG₄₀-₄₈ to study relevant CD4⁺ T
cells. As shown in Figure 5A, B6 mice that were not immunized with MOG_{40-48} contained about 250 CD4^{low} naïve CD4^{+} T cells (Fig. 5A and B), which was at the high end of the range for I-A^{b}-bound foreign peptides (Figure 1B). Recall of MOG_{40-48}-primed T cells with MOG_{40-48} peptides containing single alanine substitutions showed that the T cells depended primarily on P1, 2, 5, 7, and 8 (Figure 5C). Alanine substitution at I-A^{b} anchors at P1, 4, 6, or 7 resulted in an average reduction in T cell reactivity of 60%, which was consistent with a moderate degree of negative selection based on the results in Figure 4E. In addition, the mouse proteome only contains 19 predicted I-A^{b}-binding peptides that share 4 of 5 amino acids with MOG_{40-48} at P2, 3, 5, 7, or 8, suggesting that negative selection by homologous self peptides was probably minimal. Injection of MOG_{40-48} in CFA plus pertussis toxin resulted in the expansion of MOG_{40-48}:I-A^{b} tetramer-binding T cells, providing further evidence that not all cells in this repertoire were negatively selected. We then search bacterial proteomes for peptides that were predicted to bind I-A^{b} and matched MOG_{40-48} at P1, 2, 5, 7, and 8 but differed at P3, 4, 6, and 9. Three such peptides were identified and used to immunize B6 mice. As shown in Figure 5D and E, all three peptides induced expansion of CD4^{high} MOG_{40-48}:I-A^{b} tetramer-binding T cells to varying degrees, although not as well as MOG_{40-48} itself. Two of the three bacterial peptides, the two that stimulated the most expansion of MOG_{40-48}:I-A^{b}-specific T cells, also caused EAE, although again not as well as MOG_{40-48} (Figure 5F). These results demonstrate that a tissue-restricted self peptide that is capable of inducing autoimmunity after priming is recognized by a large naïve population that can also be activated by bacterial peptide homologs with the same TCR contact amino acids.
3.3 Discussion

This study adds to recent advances in the understanding of TCR cross-reactivity. Work from several laboratories on multiple MHCII allelomorphs has come to the conclusion that many MHCII-restricted TCRs can recognize certain peptides sharing most of the amino acids that protrude from the binding groove but differing dramatically at the amino acids that anchor the peptide to MHCII. Our results confirm this conclusion for the I-A\textsuperscript{b} MHCII molecule by showing that I-A\textsuperscript{b}-binding peptides sharing 4 of 5 putative TCR contact amino acids but no MHCII anchor amino acids with the parent peptide were immunogenic for T cells specific for parent peptide. In addition, we found that this type of TCR cross-reactivity extends to negative selection as evidenced by the 60\% reduction in the 2W109:I-A\textsuperscript{b}-specific naïve cell population in mice expressing the 2W peptide that shares only TCR contact amino acids with 2W109 peptide. These results suggest that many MHCII-restricted TCRs focus primarily on the TCR contact amino acids no matter how the peptide is anchored to MHCII. This conclusion is supported by structural biology studies showing that peptides that bind MHCII with different amino acids but have the same up pointing amino acids orient those amino acids in very similar (but non-identical) ways, and are bound with the same footprint by cross-reactive TCRs.

Many TCRs could, however, distinguish between the subtle orientations of TCR contact amino acids in peptides that share these residues but have different MHCII anchors. This conclusion is supported by the finding that some I-A\textsuperscript{b}-binding variants
sharing 4 of 5 putative TCR contact amino acids with the parent foreign peptide but with different I-A\textsuperscript{b} anchor amino acids were not immunogenic for T cells specific for the parent ligand. It is possible that the I-A\textsuperscript{b} anchor amino acids in these variant peptides alter the conformation of the TCR contact amino acids to the point of being unrecognizable to TCRs focused on the conformation of the parent peptide. The observation that the 2W109:I-A\textsuperscript{b}-specific T cells that could not recognize the related 2W:I-A\textsuperscript{b} ligand were unusually sensitive to substitutions at MHCII anchor positions in the 2W109 peptide is consistent with this possibility. It is unlikely that these TCRs directly recognized the I-A\textsuperscript{b} anchor amino acids themselves because structural studies indicate that the side chains of these residues are completely buried in the I-A\textsuperscript{b} molecule. Rather, these MHCII anchor amino acids probably created alterations in the geometry of nearby TCR contact amino acids, which can be detected by certain TCRs.

Like the 2W109:I-A\textsuperscript{b}-specific cells that survived deletion in mice expressing 2W peptide, the T cells in small naïve populations were sensitive to substitutions at both TCR contact and MHCII anchor positions in the cognate foreign peptide, while T cells in large naïve populations were usually sensitive only to substitutions in TCR contact residues. These results are consistent with the possibility that the small naïve populations are small because they experienced more negative selection than the large populations, and support earlier work by Huseby et al. showing that T cells that develop under conditions of minimal negative selection have highly cross-reactive TCRs (Huseby, White et al. 2005).

Self peptides were the likely causes of negative selection that produced small naïve populations. This contention is based on an inverse correlation between the number
of cells in a given foreign (non-mouse) peptide-specific naïve cell population and the number of predicted I-A\(^b\)-binding mouse peptides with the same TCR contacts as the foreign peptide. Although none of the 21 I-A\(^b\)-binding foreign peptides had an identical nonamer in the mouse proteome, all 21 peptides had at least 3 and as many as 181 predicted I-A\(^b\)-binding mouse peptides with 4 of 5 of the same TCR contacts as the foreign peptide. Our finding with the 2W109/2W pair showed that this degree of similarity could cause deletion of 60% of the cells in a foreign peptide-specific repertoire in a case where the self peptide was abundantly expressed in the thymus. The following hypothetical scenario illustrates how the process may work. Consider a case of 2 mouse nonamer peptides (S1 and S2) that have the same amino acids at P2, 5, 7, and 8 but different amino acids at P1, 4, and 6. Both peptides are displayed on I-A\(^b\) molecules in the thymuses of B6 mice. A third non-mouse peptide (F1) has the same amino acids at P2, 5, 7, and 8 as the two mouse peptides but different I-A\(^b\)-binding amino acids at P1, 4, and 6. The TCR contact amino acids in the 3 complexes adopt slightly different conformations, s1, s2, and f1, respectively, due to subtle effects of the different I-A\(^b\) anchor amino acids. B6 mice produce thymocyte populations with TCRs that bind s1 only, s2 only, f1 only, s1 and s2, s1 and f1, s2 and f1, and s1 and s2 and f1. Negative selection will delete all the populations with TCRs specific for s1 and/or s2, leaving a subpopulation with TCRs that bind f1 only. Since the TCRs in this population are highly specific for the f1 conformation, substitutions in any of the I-A\(^b\) anchor amino acids in the F1 peptide are likely to cause loss of TCR binding. In a different mouse strain that expresses I-A\(^b\) but lacks the S1 and S2 peptides, the F1-specific population would be
much larger, and contain cells that are not so fussy about which I-A\(^b\) anchor amino acids neighbor the TCR contact amino acids. An implication of this model is that large naïve populations are not only more helpful to the host during infection by generating large numbers of effector cells, but also because these cells express TCRs that are more tolerant to antigen variants produced by the pathogen.

This model may also be relevant for certain forms for autoimmunity caused by peptides from proteins with tissue-restricted patterns of expression. We found that one such peptide, MOG\(_{40-48}\) derived from the brain protein MOG, was recognized by a surprisingly large naïve T cell population as if had not been subjected to negative selection. This conclusion was supported by the finding that MOG\(_{40-48}:I\text{-}A^b\)-specific T cells were not sensitive to alanine substitutions at P4 or P6 of the peptide, like foreign p:I-A\(^b\)-specific T cell populations with few homologous self peptides. The MOG\(_{40-48}:I\text{-}A^b\)-specific T cell repertoire may experience minimal negative selection because relevant p: I-A\(^b\) ligands are not displayed in the thymus (Ben-Nun, Kerlero de Rosbo et al. 2006). MOG is expressed primarily in the brain so thymic antigen-presenting cells may not display MOG\(_{40-48}:I\text{-}A^b\) complexes. In addition, we found very few predicted I-A\(^b\)-binding peptides in the mouse proteome with the same TCR contact amino acids as MOG\(_{40-48}\) that could cause negative selection of MOG\(_{40-48}:I\text{-}A^b\)-specific T cells. The large size of this population may also make it susceptible to activation by I-A\(^b\)-binding microbial peptides with the same TCR contact amino acids. Activation of even of a subset of the large MOG\(_{40-48}:I\text{-}A^b\)-specific T cell by one of these mimics during infection might be enough to cause autoimmunity as in our test case shown in Figure 5F. A similar situation was
recently described for humans by De la Herran-Arita et al. CD4+ T cells with TCRs specific for a peptide from a nervous system protein called hypocretin bound to the HLA-DQ0602 MHCII molecule also recognized an influenza hemagglutinin peptide:HLA-DQ0602 complex (De la Herran-Arita, Kornum et al. 2013). The peptides shared similar TCR contact amino acids at P2, 5, 7, and 8 and also had relatively conserved substitutions at P1, 4, 6, and 9 MHCII anchor amino acids. Thus, it is conceivable that influenza infection or vaccination caused the expansion of these dual-reactive T cells that entered the nervous system and caused narcolepsy.

Finally, our results suggest that large and potentially immunodominant foreign peptides can be predicted based on a paucity of tolerance-inducing self peptide homologs. The predictive capacity of this approach is far from perfect, however, as evidenced by a relatively weak correlation between the number of predicted mouse peptide homologs and foreign peptide-specific naïve T cell population size. One problem is that peptides that share TCR contact amino acids but have different MHCII anchor amino acids are often, but not always, recognized by the same TCRs. This uncertainty means that some of the mouse peptide homologs that were predicted to cause negative selection probably do not. Another reason that many of the predicted MHCII-binding mouse peptides may not cause tolerance is that the parent protein is not expressed by thymic antigen-presenting cells or is not processed correctly to release the peptide. Thus, better prediction of the size of naïve T cell populations specific for foreign peptides will likely depend on a better understanding TCR specificity and the self peptides that cause negative selection. If these
hurdles can be cleared, it may be possible to identify the most immunogenic peptides from pathogens and use that information to design effective subunit vaccines.
Table 1

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Position (P)</th>
<th>PSSM1</th>
<th>PSSM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESAT6</td>
<td>QQWNFAEGEAAESA</td>
<td>17.95</td>
<td>17.90</td>
</tr>
<tr>
<td>FliC</td>
<td>VQNRFSAIATNLGNT</td>
<td>20.34</td>
<td>10.94</td>
</tr>
<tr>
<td>STM1540</td>
<td>VYYTTYAPQAT</td>
<td>26.35</td>
<td>25.26</td>
</tr>
<tr>
<td>STM1540</td>
<td>YTTYAPQATSA</td>
<td>23.39</td>
<td>26.35</td>
</tr>
<tr>
<td>eGFP</td>
<td>HDFFKSAMPEGYVQE</td>
<td>16.61</td>
<td>15.03</td>
</tr>
<tr>
<td>OVA3C</td>
<td>GHAHAEINEA</td>
<td>16.64</td>
<td>16.94</td>
</tr>
<tr>
<td>RpIF</td>
<td>FVSPAHII</td>
<td>29.40</td>
<td>10.03</td>
</tr>
<tr>
<td>OVA2C</td>
<td>QAVHAAHAEIN</td>
<td>10.09</td>
<td>18.32</td>
</tr>
<tr>
<td>3K</td>
<td>EAQKAKANKAVDKA</td>
<td>18.13</td>
<td>8.74</td>
</tr>
<tr>
<td>Calnex</td>
<td>LVVKNPAAHAIS</td>
<td>24.44</td>
<td>10.65</td>
</tr>
<tr>
<td>Aasf</td>
<td>VSSPAVqes</td>
<td>22.15</td>
<td>8.15</td>
</tr>
<tr>
<td>1G1W</td>
<td>EAGGALANWAVDSA</td>
<td>18.71</td>
<td>12.69</td>
</tr>
<tr>
<td>LLO</td>
<td>NEKYQAYPNVS</td>
<td>32.14</td>
<td>15.84</td>
</tr>
<tr>
<td>Derp1</td>
<td>CQIYPPPNVKI</td>
<td>22.75</td>
<td>13.08</td>
</tr>
<tr>
<td>MP98</td>
<td>HQYMTALSNEVVVF</td>
<td>29.34</td>
<td>11.81</td>
</tr>
<tr>
<td>PmpG-1</td>
<td>YVDPAAAGG</td>
<td>36.13</td>
<td>13.93</td>
</tr>
<tr>
<td>CTB</td>
<td>NNKTPHAIAAIS</td>
<td>21.57</td>
<td>13.61</td>
</tr>
<tr>
<td>GP66</td>
<td>DIYKGVQPFSSV</td>
<td>21.16</td>
<td>20.22</td>
</tr>
<tr>
<td>2W</td>
<td>EAWGALANWAVDSA</td>
<td>19.59</td>
<td>12.69</td>
</tr>
<tr>
<td>CD4Ag28m</td>
<td>VEIHRPVPGTA</td>
<td>23.68</td>
<td>14.33</td>
</tr>
</tbody>
</table>

Table 1. Nonamer cores of 20 foreign I-A<sup>b</sup>-binding peptides
Table 2

A

| Position | A | C | D | E | F | G | H | I | K | L | M | N | P | Q | R | S | T | V | W | Y |
| P1       | 13| 0 | 16| 15| 5 | 8 | 11| 7 | 7 | 10| 3 | 13| 6 | 12| 13| 20| 20| 20| 3 | 4 |
| P2       | 13| 1 | 11| 11| 6 | 29| 1  | 2 | 2 | 6 | 3 | 12| 13| 8 | 113| 1  | 1 | 24| 5 | 3 |
| P3       | 23| 1 | 6 | 12| 4 | 10| 9 | 13| 9 | 3 | 9 | 216| 7 | 23| 20| 25| 3 | 3 |
| P4       | 55| 0 | 2 | 1 | 2 | 13| 0 | 22| 1 | 3 | 1 | 3 | 54| 2 | 0 | 10| 11| 26| 0 |
| P5       | 26| 0 | 2 | 11| 0 | 18| 13| 5 | 1 | 23| 1 | 41| 1 | 11| 2 | 25| 12| 9 | 3 |
| P6       | 11| 0 | 6 | 17| 5 | 12| 2 | 18| 9 | 17| 1 | 4 | 26| 11| 10| 17| 14| 16| 1 |
| P7       | 35| 1 | 11| 11| 3 | 11| 5 | 9 | 0 | 14| 2 | 10| 4 | 4 | 34| 9 | 34| 0 |

Predicted #

14 5 10 14 9 13 5 9 12 21 5 7 13 10 12 17 11 12 2 6

B

| Position | A | C | D | E | F | G | H | I | K | L | M | N | P | Q | R | S | T | V | W | Y |
| P1       | 0.6| 0.2| 0.2| 0.1| 5.1| 0.1| 1.0| 0.6| 0.1| 0.4| 0.6| 0.4| 0.1| 0.2| 0.0| 0.1| 0.3| 1.5| 6.0| 14.7|
| P2       | 0.9| 0.0| 1.4| 1.1| 0.6| 0.6| 2.2| 0.8| 0.6| 0.5| 0.6| 1.9| 0.5| 1.2| 1.1| 1.2| 1.8| 1.5| 1.5| 0.7|
| P3       | 2.7| 0.2| 1.1| 0.8| 0.8| 2.2| 0.2| 0.2| 0.3| 0.6| 1.7| 1.0| 0.8| 0.1| 1.9| 1.4| 0.6| 2.0| 0.3|
| P4       | 2.0| 0.0| 0.1| 0.0| 0.1| 0.5| 0.2| 0.2| 0.3| 0.1| 0.0| 0.1| 9.6| 0.1| 0.1| 1.4| 0.5| 0.2| 0.0| 0.0|
| P5       | 1.6| 0.2| 0.6| 0.9| 0.5| 1.8| 0.9| 1.1| 0.4| 0.6| 1.3| 0.2| 1.6| 0.6| 1.4| 1.8| 1.9| 1.5| 0.5|
| P6       | 3.9| 0.0| 0.2| 0.1| 0.3| 1.0| 0.0| 2.4| 0.1| 0.1| 0.2| 0.4| 4.2| 0.2| 0.6| 0.6| 1.0| 2.0| 0.0| 0.0|
| P7       | 1.9| 0.0| 0.2| 0.8| 0.0| 1.4| 2.6| 0.6| 0.1| 1.1| 0.2| 5.9| 0.1| 1.1| 0.2| 1.5| 1.1| 0.7| 1.5| 0.3|
| P8       | 0.8| 0.0| 0.6| 1.2| 0.6| 0.9| 0.4| 2.0| 0.8| 0.8| 0.2| 0.6| 2.0| 1.1| 0.8| 1.0| 1.3| 1.2| 0.5| 1.5|
| P9       | 2.5| 0.2| 1.1| 0.8| 0.4| 0.8| 1.0| 1.0| 0.6| 0.7| 0.4| 1.4| 0.3| 0.4| 0.5| 2.0| 0.8| 2.6| 0.8| 0.5|

Table 2. MHCII anchoring characteristics for p:1-A\(^b\)

(A) I-A\(^b\) binding matrix produced by aligning sequences of peptides eluted from I-A\(^b\) in mass spectrometry studies.

(B) Frequency of amino acids at indicated positions, relative to expected occurrence based upon random distribution of amino acids in mouse proteome (shown in Table 1A, bottom row).
Figure 1

A

B

Unimmunized

D14 p/CFA

C

D

Log10(effector cell number)

Log10(naive cell number)

p < 0.0001

r² = 0.64
Figure 1. Relationship between p:I-A\textsuperscript{b}-specific CD4\textsuperscript{+} T cell population sizes in unimmunized and immunized mice

(A) Gating strategy for detecting p:I-A\textsuperscript{b} tetramer bound populations in secondary lymphoid tissues.

(B) Plots of double tetramer staining (far left) and CD44 expression (left) for CD4\textsuperscript{+} T cells from spleen and LN of individual naive B6 mice following enrichment with the indicated p:I-A\textsuperscript{b} tetramers labeled with two different fluorochromes. The plots on the right represent staining of the same populations in spleen and draining lymph nodes 14 days after priming subcutaneously with 10 µg of the indicated peptide emulsified in CFA.

(C) Total number of tetramer-binding CD4\textsuperscript{+} T cells for the 20 unique foreign-p:I-A\textsuperscript{b} ligands in individual unimmunized (open circles) and immunized (closed circles) mice as identified in (B) (horizontal bars indicate mean values).

(D) Linear correlation between the average Log10 (naive cell number) and Log10 (effector cell number) for each of the populations identified in (C).
Figure 2

A

![Graph showing % TCR response for samples P-1 to P10]

B

![Color Key and Heatmap with % Cross-Reactivity for various proteins]

41
Figure 2. Alanine substitution analysis of CD4$^{+}$ T cell responses to known foreign epitopes

(A) P-1 to P10 alanine substituted variant peptides were tested as stimulators of a secondary in vitro response of peptide-primed CD4$^{+}$ T cells from draining inguinal and para-aortic lymph nodes, as described in Materials and Methods. The average response at each position across 13 scanned peptides was determined as “% TCR response” as compared to the priming peptide control (100%) (+/- SD, n = 3 mice, individual wells in triplicate).

(B) Unsupervised clustering of P-1 to P10 for “% TCR response” after stimulation with alanine substituted peptides for each of the 13 indicated priming peptides.
Figure 3. Conservation of TCR contact amino acids predicts p:I-A\(^{b}\) cross-reactivity

(A) Representative plots showing staining with the indicated parent p:I-Ab tetramer of spleen and lymph node cells from mice primed with CFA alone (no peptide, negative control), bacterial-encoded homologous peptide (variant peptide), or parental peptide (positive control), 11 days after injection of subcutaneous injection.
(B) Average number of 3K:I-A<sup>b+</sup>, LLO:I-A<sup>b+</sup>, GP66:I-A<sup>b+</sup>, and 2W:I-A<sup>b+</sup> CD<sup>4+</sup> T cells after priming with the indicated peptides emulsified in CFA (+/- SEM, n = 3).

Figure 4
Figure 4. Cross-reactivity with self-p:MHCII affects foreign-p:MHCII cell number

(A) Sequence alignment of 2W and 2W109 nonamer peptides. Putative TCR contact residues shared between the peptides are in bold.

(B) Number of 2W109:I-A^b+ and LLO:I-A^b+ CD4^+ T cells in unimmunized B6 and Act-2W mice (**p < 0.0001, n.s. = not significant).

(C) P1 to P9 alanine substituted 2W109 variant peptides were tested as stimulators of a secondary in vitro response of 2W109 peptide-primed CD4^+ T cells from draining inguinal and para-aortic lymph nodes of B6 (black bars) or Act-2W (white bars) mice. The average % response, relative to 2W109 peptide stimulation (100%) is shown (+/- SD, n = 3 mice, individual wells in triplicate).

(D) Linear correlation between the average effect from each Ala-substituted variant (% reduction relative to priming peptide control) at the critical I-A^b anchor residues at P1, P4, P6, and P7 versus the Log10 (average naïve cell number).

(E) Linear correlation between the number of mouse self peptides predicted to bind I-A^b and match each of the foreign peptides in Table 1 at at least 4 TCR contact positions (P2, P3, P5, P7, P8) versus the Log10 (average naïve cell number).
Figure 5. TCR cross-reactivity with foreign peptides similar to MOG:I-A\textsuperscript{b} can elicit EAE.

(A) Plots of double tetramer staining (top) and CD44 expression (bottom) for MOG:I-A\textsuperscript{b}-specific CD4\textsuperscript{+} T cells from spleen and LN of individual unimmunized B6 mice following enrichment with MOG:I-A\textsuperscript{b} tetramer labeled cells.

(B) Total number of MOG:I-A\textsuperscript{b} tetramer-binding CD4\textsuperscript{+} T cells in individual unimmunized mice (horizontal bars indicate mean values).

(C) P1 to P9 alanine substituted MOG variant peptides were tested as stimulators of a secondary in vitro response of MOG peptide-primed CD4\textsuperscript{+} T cells from draining inguinal and para-aortic lymph nodes of B6 mice. The average % response, relative to MOG peptide stimulation (100%) is shown (+/- SD, n = 3 mice, individual wells in triplicate).

(D) Representative plots indicating clonal expansion and CD44 expression of MOG:I-A\textsuperscript{b}-specific cells 14 days after priming subcutaneously with 10 \(\mu\)g of the indicated homologous foreign peptide emulsified in CFA.

(E) Average number of MOG:I-A\textsuperscript{b}+ CD4\textsuperscript{+} T cells after priming with 10 \(\mu\)g of the indicated peptides (+/- SEM, n = 3).

(F) EAE disease course after subcutaneous immunization with the 100 \(\mu\)g of the indicated peptides in (E) plus intravenous injection of 200 ng pertussis toxin on day 0 and day 2, as described in Materials and Methods.
Chapter 4

CD4+ T cell persistence and function after infection is maintained by low-level peptide:MHCII presentation

4.1 Introduction

Expanded populations of antigen-specific T and B cells that persist long after pathogen clearance are responsible for immunological memory. These cells are useful to the host, providing protective immunity to subsequent challenge by the original microbe. The memory cell paradigm has been established most clearly for CD8+ T cells (Lau, Jamieson et al. 1994). CD8+ T cells expressing TCRs specific for microbe peptide:MHCI ligands proliferate extensively, producing a peak number of effector cells about a week after acute infection. About 90% of these effector cells then die by apoptosis, leaving a population of memory cells that is stably maintained by recurrent IL-15-driven homeostatic proliferation. Central memory CD8+ T cells (Tcm) re-circulate through secondary lymphoid organs, producing new memory cells during secondary responses, while effector memory (Tem) cells located in non-lymphoid organs are immediately cytotoxic during secondary responses.

It is less clear whether the concept of stable immune memory applies to CD4+ T cells. Naïve CD4+ T cells expressing TCRs specific for microbe peptide:MHCII ligands proliferate extensively to produce a peak number of effector cells about a week after acute infections with Lymphocytic Choriomeningitis Virus (LCMV) or Listeria
monocytogenes (Homann, Teyton et al. 2001, Pepper, Linehan et al. 2010). As in the case of CD8<sup>+</sup> T cells, about 90% the effector cells then die, leaving a population of memory-phenotype cells, about half of which are Th1 cells and the other half follicular helper cell-like Tcm cells (Pepper, Pagán et al. 2011). Although both CD4<sup>+</sup> and CD8<sup>+</sup> T cells are maintained by recurrent IL-15-driven homeostatic proliferation, the proliferation rate is much lower for CD4<sup>+</sup> memory T cells than for CD8<sup>+</sup> memory T cells (Sprent 2008, Pepper, Linehan et al. 2010). Thus, CD4<sup>+</sup> memory T cells slowly decline in number after infection is cleared, probably because their lower rate of homeostatic proliferation cannot keep pace with their death rate.

Previous studies showing the decline of CD4<sup>+</sup> memory T cells involved acute infections, which were completely cleared by CD8<sup>+</sup> T cells and innate immune cells with no apparent involvement of CD4<sup>+</sup> T cells (Matloubian, Concepcion et al. 1994, Portnoy, Auerbuch et al. 2002). Infections that are controlled by CD4<sup>+</sup> T cells (Reith and Mach 2001) are caused by microbes such as Mycobacterium, Leishmania, and Salmonella species, that establish persistent infections in the phagosomes of phagocytes. In these cases, CD4<sup>+</sup> T cells contain, but never fully clear the microbes from the initial site of infection, while providing sterilizing immunity at all other body sites (Belkaid, Hoffmann et al. 2001, Monack, Mueller et al. 2004, Urdahl, Shafiani et al. 2011). Notably, clearance of the original infection is associated with loss of systemic immunity (Belkaid, Piccirillo et al. 2002), indicating that the maintenance of protective CD4<sup>+</sup> T cells depends on persistent antigen, as suggested by several recent studies (Lin, Kemball et al. 2010, Reiley, Shafiani et al. 2010, Stephens and Langhorne 2010). The requirement for
persistent antigen presentation for CD4\(^+\) T cell maintenance goes against the idea derived from studies of CD8\(^+\) T cells that durable memory and immune protection develop after antigen is cleared while exhaustion ensues if it is not cleared.

Here, we assessed CD4\(^+\) T cell responses to the facultative intracellular bacterium *Salmonella enterica* that can infect mice and humans through the gastrointestinal tract. Mice expressing Nramp1, a protein that helps limit bacterial replication in phagocytes, develop a persistent infection following oral inoculation, which is controlled by IFN-\(\gamma\)-producing CD4\(^+\) T cells (de Jong, Altare et al. 1998, Monack, Bouley et al. 2004). In the current study, we developed a model to study a p:MHCII-specific immune response to *Salmonella enterica* serovar Typhimurium (ST) infection in Nramp1-resistant (Nramp1\(^{Gly169/Gly169}\), referred to hereafter as Nramp1\(^R\)) mice to understand how bacterial persistence and p:MHCII presentation affect the numerical and functional stability of protective CD4\(^+\) T cells.

### 4.2 Results

**Generation of Salmonella expressing an immunodominant CD4\(^+\) T cell epitope**

Before monitoring the stability of CD4\(^+\) T cells during persistent ST infection, it was necessary to identity a relevant epitope. Although epitopes consisting of I-A\(^b\) molecules complexed with peptides from the FliC and SseJ proteins of ST have been described (Lee, McLachlan et al. 2012), the CD4\(^+\) T cell populations that recognize them are very small (Moon, Chu et al. 2007). One of the largest naïve CD4\(^+\) T cell populations identified to date is specific for a peptide called 2W (Moon, Chu et al. 2007), which is a
variant of MHCII I-E alpha chain peptide 52–68 (Rees, Bender et al. 1999). We therefore used the lambda Red-mediated recombination system to insert the 2W peptide coding sequence at the 3’ end of the OmpC gene in the ST chromosome (Uzzau, Figueroa-Bossi et al. 2001). The resulting organisms, referred to hereafter as ST-2W, express an OmpC-2W fusion protein (Fig. 6A). As shown in Fig. 6B, oral inoculation of ST-2W bacteria into Nramp1R 129 mice resulted in an infection that peaked between 2-3 weeks in the mesenteric lymph nodes (MLN) and spleen, then declined to an undetectable level in the spleen by day 50 but persisted at low levels in MLN for hundreds of days as previously described by Monack and colleagues for wild-type ST organisms (Monack, Bouley et al. 2004). Therefore, ST-2W organisms were capable of producing a persistent infection that was controlled by CD4+ T cells.

The CD4+ T cell response to intragastric ST-2W infection

We next measured the number of 2W:I-A\(^b\)-specific cells over time after infection using a p:MHCII tetramer-based cell enrichment approach (Moon, Chu et al. 2009). Cells from MLN and spleens of infected 129 mice were stained with fluorochrome-labeled 2W:I-A\(^b\) tetramer and anti-fluorochrome-labeled magnetic beads and enriched on magnetized columns as previously described (Moon, Chu et al. 2007). 2W:I-A\(^b\) tetramer-binding cells were detected by flow cytometry among the CD4+ cells that bound to the column.

Uninfected 129 mice each contained about 100 2W:I-A\(^b\)-specific CD4+ T cells in MLN and about 300 in spleen. The majority of the cells in both locations were CD44\(^{low}\)
as expected for naïve T cells (Fig. 7A). A similar number and phenotype of 2W:I-A\textsuperscript{b}-specific CD4\textsuperscript{+} T cells was observed in these tissues in 129 mice infected with ST organisms, not expressing the 2W peptide (Fig. 7A). In contrast, 2W:I-A\textsuperscript{b}-specific CD4\textsuperscript{+} T cells increased dramatically in number and CD44 expression in MLN and spleen after ST-2W infection (Fig. 7A). 2W:I-A\textsuperscript{b}-specific T cells present on day 6 and 60 after ST-2W infection expressed large amounts of CD44 (Fig. 7A) and the Th1 lineage-defining transcription factor T-bet (Fig. 7B). Therefore, intragastric ST-2W infection of Nramp\textsuperscript{R} mice induced an early homogeneous Th1 population, some of which survived into the persistent phase of the infection.

**Salmonella p:MHCII-specific CD4\textsuperscript{+} T cells are numerically stable during persistent infection**

We next studied the kinetics of 2W:I-A\textsuperscript{b}-specific T cells following intragastric infection of 129 mice to determine whether a stable cell population was produced. The naïve 2W:I-A\textsuperscript{b}-specific CD4\textsuperscript{+} T cell population of about 400 cells proliferated to produce 3x10\textsuperscript{5} cells in MLN and 2x10\textsuperscript{6} cells in the spleen two weeks after infection (Fig. 7C). The number of cells in both locations then underwent an 80% reduction that stabilized around day 40. Thereafter, the number of the 2W:I-A\textsuperscript{b}-specific CD4\textsuperscript{+} T cells remained consistent at 3x10\textsuperscript{4} cells in MLN and 3x10\textsuperscript{5} cells in the spleen. The cells in both locations were CD44\textsuperscript{high} T-bet\textsuperscript{high} cells (Fig. 7B). These results show that ST p:MHCII-specific memory-phenotype cells are stably maintained during persistent ST infection.
A BrdU-labeling experiment was then performed to determine whether proliferation within chronically infected tissues could account for CD4+ T cell stability during the persistent phase of infection. 129 mice were injected with BrdU on day 50 after ST-2W infection, when viable bacteria were detected only in the MLN (Fig. 6B), and lymphoid organs were analyzed for BrdU incorporation 12 hours later as a measure of cellular proliferation. This short labeling period was chosen to maximize the chance that cells were labeled in the location where they divided. As shown in Fig. 8B, 5-10% of the 2W:I-Ab-specific CD4+ T cells located in the MLN labeled with BrdU, whereas only 2-3% of cells in the spleen were labeled. Therefore, the stability of the 2W:I-Ab-specific population correlated with a higher rate of proliferation in the persistently infected MLN.

**CD4+ T cell maintenance is dependent upon persistent p:MHCII presentation**

If the 2W:I-Ab-specific population was maintained by proliferation of a subset of cells in persistently infected MLN, then clearance of the ST-2W bacteria would be expected to result in decline of the population. ST-2W-infected mice were cleared of infection by treatment with enrofloxacin antibiotic for 40 days, beginning at day 28 after ST-2W infection, as a test of this hypothesis. As shown in Fig. 9B, mice not treated with antibiotic had about twice as many 2W:I-Ab-specific T cells in their lymphoid organs on day 68 of ST-2W infection as mice that were treated with antibiotic.

We next determined whether p:MHCII recognition was the aspect of persistent infection that was required for the stability of the 2W:I-Ab-specific population. CFSE labeled 2W:I-Ab-specific cells from the spleens and lymph nodes of persistently infected
(CD45.1 x 129) F1 mice were transferred into (B6 x 129) F1 mice with time-matched ST-2W or wild-type ST infections. The number of donor-derived CD45.1+ 2W:I-Ab+ CD4+ T cells was enumerated 1, 7, or 50 days after transfer to assess the stability of the transferred cells. ST-2W bacteria were not transferred along with the T cells, as evidenced by the fact that the 2W:I-Ab-specific cells of endogenous origin in wild-type ST-infected mice, receiving 2W:I-Ab-specific memory-phenotype cells, neither proliferated nor increased CD44 expression (Fig. 9C). As shown in Fig. 9D, the numbers of transferred 2W:I-Ab-specific cells were similar between days 7 and 50 in ST-2W-infected hosts, but were significantly lower on day 50 than on day 7 in hosts infected with wild-type ST bacteria lacking the 2W peptide.

The stability of 2W:I-Ab-specific cells in ST-2W-infected hosts was associated with proliferation. 2W:I-Ab-specific cells residing in ST-2W- or wild-type ST-infected hosts for 50 days underwent 1-3 divisions, probably in response to IL-15 (Pepper, Linehan et al. 2010). However, donor-derived 2W:I-Ab-specific cells that resided in ST-2W infected hosts for 50 days also contained cells that had divided greater than 7 times. Even though these cells accounted for 70% of the population 50 days after transfer, the fact they doubled at least 7 times during that period means they arose from less than 1% of the initial transferred population. These results are consistent with the possibility that periodic proliferation by a small number of cells in persistently infected MLN accounted for the long-term numerical stability of the 2W:I-Ab-specific cell population.

**ST-2W induced Th1 cells are multifunctional**
We then investigated whether persistent ST-2W infection affected the function of 2W:I-A\textsuperscript{b}-specific T cells. Mice infected with ST-2W 14, 45, or 100 days earlier were injected intravenously with 100 µg of 2W peptide to elicit lymphokine production by 2W:I-A\textsuperscript{b}-specific T cells, which was detected by direct ex vivo intracellular staining. 2W:I-A\textsuperscript{b}-specific T cells increased CD69 following 2W peptide injection and produced IFN-γ as expected based on their high expression of T-bet. While IFN-γ is essential for controlling persistent ST infection (Monack, Bouley et al. 2004), the best correlates of a protective Th1 response for other infections are CD4\textsuperscript{+} T cells capable of simultaneously producing IFN-γ, TNF, and IL-2 (triple\textsuperscript{+}) (Darrah, Patel et al. 2007, Lindenstrom, Agger et al. 2009). Interestingly, almost all of the 2W:I-A\textsuperscript{b}-specific cells also expressed TNF and a fraction of these cells also produced IL-2 (Fig. 10A). Thus, 2W:I-A\textsuperscript{b}-specific T cells were maintained in a highly multifunctional state during persistent ST-2W infection.

**Antigen abundance affects CD4\textsuperscript{+} T cell function**

We next tested whether the low-level, localized p:MHCII presentation in the MLN of ST-2W infected mice that maintains CD4\textsuperscript{+} T cell numbers also preserved their function. We found that the overall composition of single (IFN-γ\textsuperscript{+}), double (IFN-γ\textsuperscript{+}, TNF\textsuperscript{+}) and triple lymphokine-producing (IFN-γ\textsuperscript{+}, TNF\textsuperscript{+}, and IL-2\textsuperscript{+}) cells was similar between donor-derived cells that resided for 50 days in ST-2W infected or uninfected hosts (Fig. 11C). These results indicate that periodic stimulation by 2W:I-A\textsuperscript{b} complexes during the persistent phase of infection did not significantly affect the cytokine production potential of the 2W:I-A\textsuperscript{b}-specific cells.
This result was surprising in light of studies indicating persistent stimulation by p:MHCI complexes results in functional exhaustion of CD8+ T cells (Virgin, Wherry et al. 2009). It was possible that 2W:I-A\textsuperscript{b}-specific cells avoided exhaustion in mice with persistent ST-2W infection because infected antigen presenting cells were too rare. We tested this possibility by assessing the fate of 2W:I-A\textsuperscript{b}-specific memory-phenotype cells after transfer into (Act-2W x 129) F\textsubscript{1} transgenic mice that display 2W:I-A\textsuperscript{b} throughout the body due to ubiquitous expression of 2W peptide under the control of the Actb promoter (Moon, Dash et al. 2011). Indeed, 2W:I-A\textsuperscript{b}-specific cells induced by ST-2W infection completely lost the capacity to produce lymphokines in response to peptide challenge after transfer and residence for 50 days in ST-2W-infected (Act-2W x 129) F\textsubscript{1} transgenic mice. This loss of function was accompanied by induction of surface PD-1, which is a marker of T cell exhaustion (Fig. 11A) (Barber, Wherry et al. 2005). Therefore, 2W:I-A\textsuperscript{b}-specific cells became functionally exhausted when exposed to systemic 2W:I-A\textsuperscript{b} complexes.

4.3 Discussion

Previous studies of acute \textit{L. monocytogenes} and LCMV infections reached the conclusion that p:MHCII-specific CD4\textsuperscript{+} memory-phenotype T cells slowly decline after infections are cleared (Homann, Teyton et al. 2001, Pepper, Linehan et al. 2010). The numerical decline of murine CD4\textsuperscript{+} memory-phenotype T cells without their relevant antigen is at odds with the remarkable stability of CD8\textsuperscript{+} memory-phenotype T cells, IgM\textsuperscript{+} memory B cells and plasma cells (Slifka, Antia et al. 1998, Homann, Teyton et al. 2001,
Pape, Taylor et al. 2011), and challenges the classical definition of immune memory. The stability of the population of 2W.1-A\textsuperscript{b}-specific T cells in mice with persistent ST-2W infection could result from the periodic proliferation of a few memory-phenotype CD4\textsuperscript{T} cells to produce short-lived effector cells coupled with the slow death of non-stimulated memory-phenotype cells. Numerical stability would be achieved in this case despite the fact that neither the short-lived effector cells nor the quiescent memory-phenotype cells are perfectly stable. A similar scenario has been reported for CD8\textsuperscript{T} cells induced by chronic LCMV infection (Shin, Blackburn et al. 2007).

The numerical decline of CD4\textsuperscript{T} memory-phenotype T cells without p:MHCII presentation may be less surprising when considering the relationship between immune memory and protection. Immune memory is useful to the host only insofar as it contributes to immune protection. CD4\textsuperscript{T} T cells are not essential for immunity to acute L. monocytogenes and LCMV infections, which are controlled by CD8\textsuperscript{T} T cells specific for abundant p:MHC\textsubscript{I} ligands generated by these cytosolic intracellular microbes (Matloubian, Concepcion et al. 1994, Portnoy, Auerbuch et al. 2002). Thus, the decline of CD4\textsuperscript{T} T cells comes at no cost to the host in these cases because these cells do not provide a protective advantage. In contrast, in a case where CD4\textsuperscript{T} T cells play an essential role in controlling a persistent phagosomal ST infection, p:MHCII-specific CD4\textsuperscript{T} memory-phenotype T cells were numerically stable. The fact that this stability depended on persistent infection and p:MHCII presentation indicates TCR stimulation is the key signal needed for maintenance of CD4\textsuperscript{T} T cells. Therefore, the classical idea of memory cell stability without antigen does not have to apply to CD4\textsuperscript{T} T cells in cases
where they are protective because the phagosomal infections that they control are persistent, thus providing a “constant TCR reminder” to maintain the T cell population.

In the cases of chronic LCMV, hepatitis C, and HIV infections, prolonged p:MHCI antigen stimulation eventually results in the exhaustion of CD8⁺ T cells (Shin and Wherry 2007). The fact that persistent ST-2W infection and p:MHCII presentation was required to maintain CD4⁺ T cells without exhaustion is probably related to the low level of peptide:MHCII antigen presentation during persistent ST infection. This is supported by the recent demonstration of CD8⁺ T cell function in the face of chronic HSV infection in the sensory ganglia (Mackay, Wakim et al. 2012). The restricted anatomical location of infection could also influence CD4⁺ T cell function during persistent ST infection. After the first two months after oral ST infection, the MLN were the primary site in the body to harbor viable bacteria. Thus, it is conceivable that CD4⁺ T cells periodically migrate into sites of low-level persistent infection, proliferate in response to p:MHCII ligands displayed by persistently infected phagocytes, and then migrate to other body sites to recover. This process could produce numerical stability by balancing the death of cells in the population and prevent the cells from becoming exhausted. These ideas are supported by our observations that MLN were preferential sites of CD4⁺ T cell proliferation in persistently infected mice and exposure to systemic p:MHCII complexes resulted in CD4⁺ T cells that could not produce lymphokines and expressed high levels of PD-1, the canonical marker for exhausted cells (Shin and Wherry 2007). Thus, although the lack of cytokine production by these cells could be due
to an indirect mechanism like regulatory T cell suppression, it more likely relates to the
direct mechanism of T cell exhaustion.

Persistent low-level p:MHCII presentation by phagocytes during ST infection was
also associated with homogenous production of Th1 cells, unlike transient p:MHCII
presentation during acute infections, which generate a mixture of Th1 cells and follicular
helper cells (Marshall, Chandele et al. 2011, Pepper, Pagán et al. 2011). The reason that
follicular helper cells were not maintained during persistent ST infection could relate to
lack of p:MHCII presentation by B cells (Johnston, Poholek et al. 2009).

Given the impressive lymphokine production capacity of 2W:1-A\textsuperscript{b}-specific cells
in the face of persistent ST-2W infection, it was possible that periodic TCR stimulation
was actually required to maintain multifunctional Th1 cells with maximal IFN-\gamma
production. However, our results demonstrate that T-bet expression and IFN-\gamma production
capacity was maintained at a high level in CD4\textsuperscript{+} memory-phenotype T cells that were
parked in antigen-free hosts. However, a loss in the number of highly functional Th1 cells
would be expected to reduce the capacity of CD4\textsuperscript{+} T cells to control a subsequent
phagosomal ST infection. Such a loss could explain the concomitant immunity
phenomenon studied in Leishmania-infected individuals. In this case, IFN-\gamma producing
Th1 cells control the infection within phagocytes at the initial site of infection and
prevent it from spreading to other parts of the body. Oddly however, the Th1 cells never
eliminate the microbes from the initial site. Indeed, persistent infection at the original site
is required for the Th1 cells to eliminate bacteria from other body sites after a second
infection (Belkaid, Piccirillo et al. 2002). Our results indicate that localized persistent
infection and p:MHCII presentation is required to maintain a numerically stable CD4\(^+\) T cell population with maximal IFN-\(\gamma\) production capacity needed to control the infection at the initial site and eliminate it from other sites.

The fact that CD4\(^+\) T cells control persistent ST infection in phagocytes and presumably depend on p:MHCII presentation by these cells to be maintained as maximally functional cells and yet cannot eliminate the infection, is perplexing. One possibility is regulatory T cells within MLN restrain microbicidal activities of CD4\(^+\) effector T cells in the same location (McLachlan, Catron et al. 2009, Johanns, Ertelt et al. 2010). Alternatively, persistently infected phagocytes may produce IL-10, which can ameliorate some of the phagocyte-activating effects of IFN-\(\gamma\) made by CD4\(^+\) T cells engaging in cognate interactions.

Our results suggest that a CD4\(^+\) T cell-dependent vaccine for a phagosomal pathogen will have to produce a local and long-lasting depot of bacterial antigen to maintain the protective CD4\(^+\) T cell population, which would otherwise decline. The failure to maintain protective CD4\(^+\) T cells could explain why the effectiveness of the Bacillus Calmette-Guérin vaccine against tuberculosis wanes after the vaccine organisms are cleared from the body. To improve on this situation, it may be necessary to produce even longer-lasting antigen delivery systems, perhaps along the lines of those used for long term administration of contraceptives.
Figure 6

A

BAP  Wild-type  OmpC-2W

50 kDa

37 kDa

FLAG

MLN  Spleen

B

CFUs/organ

≤10

10

10

10

10

7-14

50

>100

Day
Figure 6. Intragastric ST-2W infection is controlled by CD4⁺ T cells and persists in the MLN

(A) Western blot of ST bacteria genomically tagged to express the 2W epitope on outer membrane porin C (OmpC) (42kDa). Wild-type ST bacteria do not express this peptide and are shown as a negative control. BAP is a FLAG-fusion protein of 47 kDa that served as a positive control. (B) ST-2W bacterial CFUs in the MLN (filled circles) or spleens (open circles) of mice with horizontal bars for mean values at the indicated times after intragastric ST-2W infection.
Figure 7

A

Uninfected  ST infected  ST-2W infected
MLN  Spleen  MLN  Spleen

CD44

0.21  1.7  73
2.2  1.1  87

B

Day 6  Day 60
MLN  Spleen  MLN  Spleen

% of Maximum

89  89
95  89

C

T-bet

% of Maximum

89  89
95  89

2W:I-A^b+ cells

Spleen  MLN

Time (d)

10^2  10^3  10^4  10^5  10^6  10^7

0  100  200  300  400  500
Figure 7. ST-2W infection induces a stable population of 2W:I-A<sup>b</sup>-specific T cells

(A) Plots of CD4<sup>+</sup> T cells from uninfected (left) or at day 21 after ST (middle) or ST-2W (right) infection in 129 mice. The percentages of 2W:I-A<sup>b+</sup> cells in the tetramer-enriched populations are shown. (B) Histograms showing T-bet on naïve CD4<sup>+</sup> CD44<sup>low</sup> cells (filled gray) or 2W:I-A<sup>b+</sup>-specific CD44<sup>high</sup> T cells (black line) from the MLN (top) and spleens (bottom) of mice at the indicated times after intragastric infection with ST-2W bacteria. (C) Mean ± SEM (n ≥ 3 for each time point) of 2W:I-A<sup>b+</sup> T cells in the spleens (filled circles) and MLN (open circles) over the first 552 days after intragastric infection with ST-2W bacteria.
Figure 8

(A) BrdU labeling of 2W:I-A^b-specific T cells in the MLN (left) and spleen (right) of mice injected 12 hours earlier with 1 mg of BrdU 50 days after intragastric infection with ST-2W bacteria. (B) Percentage of BrdU+ 2W:I-A^b-specific T cells in the MLN (filled circles) and spleens (open circles) of mice infected with ST-2W bacteria. Each symbol represents an individual mouse and horizontal line indicates the mean.

Figure 8. CD4^+ T cell proliferation occurs at the site of persistent infection
Figure 9. Numerical stability of CD4⁺ T cells is dependent on p:MHCII-induced proliferation

(A) Plots of CD4⁺ T cells at day 68 after ST-2W infection from untreated mice (top) or mice that were treated with enrofloxacin for 40 days beginning at day 28 post-infection.
(bottom). (B) Number of 2W:I-A^b-specific cells in spleen and lymph nodes at day 68 after ST-2W infection in untreated (filled circles) or enrofloxacin-treated (open circles) mice. (C) CD4^+ T cells from ST-2W immune mice were transferred into ST-2W (left) or ST (right) infection-matched hosts 50-75 days post-infection. Plots of 2W:I-A^b-enriched CD4^+ T cells 1 day (top) or 50 days (bottom) after transfer. Gates indicate donor-(CD45.1^+) or host-derived (CD45.1^-) 2W:I-A^b^+ cells. Day 50 gates indicate frequency of CFSE fully (left gate) or partially (right gate) diluted, donor-derived (CD45.1^+) 2W:I-A^b^+ cells. (D) Frequency of donor-derived 2W:I-A^b^-specific cells at 7 or 50 days after transfer into ST-2W (filled circles) or ST (open circles) infection-matched hosts, normalized for day 1 park rates in independent experiments. P values are from a comparison of day 7 and day 50 values from the ST-2W- or ST-infected recipients. (E) Number of CFSE fully diluted, donor-derived (CD45.1^+) 2W:I-A^b^-specific cells at 50 days after transfer into ST-2W (filled circles) or ST (open circles) infection matched hosts.
Figure 10

A

IL-2^ - IL-2^ +

Day 14

Day 45

Day 90

IL-2

IFN-γ

TNF

B

% of total 2Wt-A^+ cells

Day 14  Day 45  Day 90

MLN

Spleen
Figure 10. ST-2W infection induces multifunctional Th1 cells

(A) Histograms showing intracellular IL-2 (left) in 2W:I-A\(^b\)-specific cells and contour plots of TNF versus IFN-\(\gamma\) of IL-2\(^-\) (middle) and IL-2\(^+\) (right) cells from the MLN 2 hrs after intravenous injection of 2W peptide at 14 (top), 45 (middle), or 90 (bottom) days after intragastric ST-2W infection. The shaded histogram shows intracellular IL-2 in 2W:I-A\(^b\)-specific cells from mice that were not injected with 2W peptide. The day 14 and 45 experiments were collected on the flow cytometer on the same day, while the day 90 experiment was run on a different day. The quadrants were set on each day on CD44\(^\text{low}\), 2W:I-A\(^b\)- cells, which do not produce cytokines. (B) Percentages of single\(^+\) (IFN-\(\gamma\)^+), double\(^+\) (IFN-\(\gamma\)^+ , TNF\(^+\)) triple\(^+\) (IFN-\(\gamma\)^+ , TNF\(^+\), and IL-2\(^+\)), and triple\(^-\) (IFN-\(\gamma\)^-, TNF\(^-\), and IL-2\(^-\)) lymphokine-producing 2W:I-A\(^b\)-specific cells in the MLN (left) and spleens (right), 2 hrs after intravenous injection of 2W peptide at the indicated days after ST-2W infection.
Figure 11. Function of CD4$^+$ memory-phenotype T cells is dependent upon antigen abundance

(A) Plots of CD45.1 versus 2W:I-A$^b$ tetramer (left) or histograms of PD-1 expression by CD45.1$^+$ 2W:I-A$^{b^+}$ donor-derived cells (right) 50 days after transfer into ST-2W infection-matched (B6 x 129) F$^1$ hosts (top), uninfected (B6 x 129) F$^1$ hosts (middle), or ST-2W infection-matched (Act-2W x 129) F$^1$ transgenic hosts (bottom). (B) Histograms
showing intracellular IFN-γ, TNF, or IL-2 expression in donor-derived (CD45.1⁺) 2W:I-Aᵇ-specific cells 50 days after transfer into ST-2W infection-matched (B6 x 129) F₁ (black line), uninfected (B6 x 129) F₁ (dotted line), or ST-2W infection-matched (Act-2W x 129) F₁ transgenic hosts (gray), before (top) or 2 hr after (bottom) intravenous injection of 2W peptide. (C) Percentages of single⁺ (IFN-γ⁺), double⁺ (IFN-γ⁺, TNF⁺) triple⁺ (IFN-γ⁺, TNF⁺, and IL-2⁺), and triple⁻ (IFN-γ⁻, TNF⁻, and IL-2⁻) lymphokine-producing donor-derived (CD45.1⁺) 2W:I-Aᵇ-specific cells 50 days after transfer into ST-2W infection-matched (B6 x 129) F₁, uninfected (B6 x 129) F₁, or ST-2W infection-matched (Act-2W x 129) F₁ transgenic hosts, 2 hrs after intravenous injection of 2W peptide. Horizontal bars indicate mean values.
A large population of CD4$^+$ T cells specific for a single p:MHCII plays a large role in control of *Salmonella enterica* infection

## 5.1 Introduction

The relationship between naïve T cell population size and the response to immunization has brought new insight to understanding immunodominance. The important role for large naive p:MHC-I-specific CD8$^+$ T cell populations in controlling viral infections is now clear. For example, overlapping peptide libraries encompassing the entire proteome of LCMV have led to a comprehensive assessment of the LCMV-peptide:MHC-I-specific CD8$^+$ T cell responses in mice. Interestingly, of 28 identified LCMVp:D$^b$-specific CD8$^+$ T cell responses, the 3 largest naïve cell populations accounted for almost one third of the total CD8$^+$ T cell response to LCMV infection, whereas the 18 smallest naïve cell populations accounted for a very small percentage (Kotturi, Scott et al. 2008). Furthermore, vaccinating mice with individual peptides recognized by naïve CD8$^+$ T cell populations consisting of 100 cells or more protected mice against LCMV challenge (van der Most, Murali-Krishna et al. 1998, Kotturi, Peters et al. 2007, Kotturi, Scott et al. 2008).

However, whether a large population of CD4$^+$ T cells specific for a single bacterial peptide can have a meaningful impact on protecting the host from bacterial infection is unknown. *Salmonella enterica* is a facultative intracellular pathogen that is
controlled by IFN-γ-producing CD4⁺ T cells (de Jong, Altare et al. 1998, Monack, Bouley et al. 2004). In mice, S. enterica serovar Typhimurium (ST) causes a persistent infection similar to typhoid fever in humans. To determine whether a single large p:MHCII-specific CD4⁺ T cell population could influence the control of persistent ST infection, we studied the responses to infection with either wild type ST or recombinant ST expressing the 2W peptide (ST-2W). There was a significantly lower bacterial burden in ST-2W infected mice as compared to paired ST infection. By comparing this outcome to mice that could not respond to 2W:I-Aᵇ but could respond to all other ST expressed peptides, it was determined that this effect was due to the large response to 2W:I-Aᵇ.

5.2 Results

The 2W:I-Aᵇ-specific response to ST-2W infection accounts for a large fraction of the total CD4⁺ T cell response

The 2W variant of MHCII I-E alpha chain peptide 52–68 (Rees, Bender et al. 1999) is a foreign peptide in B6 and 129 mice, which binds to the MHCII molecule I-Aᵇ and is recognized by a large population of approximately 300 naïve CD4⁺ T cells. This large naïve population is in contrast to the small populations of 20 or fewer cells that recognizing known ST-encoded p:I-Aᵇ ligands FliC and SseJ (Moon, Chu et al. 2007, Lee, McLachlan et al. 2012). This prompted the question of what proportion of the protective IFN-γ-producing CD4⁺ T cell response to ST-2W infection can be attributed to 300 naïve CD4⁺ T cells specific for 2W:I-Aᵇ. Considering Salmonella expresses
approximately 4,000 foreign proteins and many would contain multiple I-A\(^b\) binding peptides, we hypothesized that 2W:I-A\(^b\)-specific response would comprise at most 1% of the total IFN-\(\gamma\) response to ST-2W despite its large naïve population. To test this, we immunized 129 mice by oral inoculation with \(10^8\) ST-2W organisms re-stimulated CD4\(^+\) T cells in the spleen and mesenteric lymph nodes with the non-specific TCR mimics PMA and ionomycin. As shown in Figure 12A, the 2W:I-A\(^b\)-specific response accounted for between 10-30 percent of the total number of IFN-\(\gamma\) producing cells.

**The response to a single p:MHCII lowers the bacterial load during persistent ST infection**

The surprising frequency of IFN-\(\gamma\) producing that were specific for a single p:MHCII ligand lent itself to the possibility that this could be affecting immunity. A clue that this may be the case was that the bacterial load detected in mice infected with ST-2W (Nelson, McLachlan et al. 2013) was lower than originally described for infection with the ST strain used to generate ST-2W (Monack, Bouley et al. 2004). It was previously assumed that the most likely explanation for this discrepancy was that insertion of the 2W peptide on the C-terminal end of the essential outer membrane protein OmpC affected the fitness of the bacteria. To test these two non-mutually exclusive possibilities, we infected (B6 x 129) F\(_1\) and \((Act-2W\) x 129) F\(_1\) mice with either ST or ST-2W and determined the number of bacterial CFUs 50 days later. The presence of 2W as a self peptide in Act-2W mice leads to complete tolerance of 2W:I-A\(^b\)-reactive cells (Moon, Dash et al. 2011). To our surprise, we found significantly higher ST-2W CFUs in \((Act-2W\) x 129) F\(_1\) mice than
(B6 x 129) F_1 mice (Figure 12B). In contrast, there was no difference between (B6 x 129) F_1 mice and (Act-2W x 129) F_1 mice in controlling the control ST infection. It should also be noted there were still more ST bacteria than ST-2W bacteria in either mouse strain, suggesting that the epitope tag also affected the survival of the bacteria. Nevertheless, this result demonstrates that a large population of CD4^+ T cells specific for a single p:MHCII plays a large role in control of Salmonella enterica infection.

5.3 Discussion

The fact that the 300 naïve cells specific for 2W:I-A^b could produce 30% of the total response to infection with a complex bacterial pathogen highlights importance of identifying large naïve CD4^+ T cell populations. This result also demonstrates how large naïve populations are relatively rare in the pre-immune repertoire, which could explain the paucity of immunodominant epitopes that have been described for these pathogens. Therefore, an improved understanding of the factors that shape naïve T cell population size is necessary to identify the key targets for subunit vaccines.
Figure 12

A

![Diagram showing CD4+ T cell population with IFN-γ and 2W:I-Ab gate]

B

![Graph showing CFUs with P < 0.05 and n.s.]

Figure 12. The response of a single large CD4+ T cell population affects the bacterial load during persistent ST infection.

(A) Plots indicating the percentage of IFN-γ-producing cells CD4+ T cells by intracellular cytokine staining after PMA and ionomycin re-stimulation (left) and downstream gate showing frequency of IFN-γ+ cells that are 2W:I-Ab+ after ST-2W infection.
(B) ST-2W bacterial CFUs (filled circles, left) or ST bacterial CFUs (open circles, right) in the MLN of (B6 x 129) F1 or (Act-2W x 129) F1 mice, 50 days after oral infection. Horizontal bars indicate mean values.
Chapter 6

Conclusions

The latter half of my thesis research, presented in chapter 3, sought to understand fundamental principles of TCR recognition of p:MHCII ligands to determine if general rules could be established to predict what peptides will elicit large CD4\(^+\) T cell responses. The result of these efforts was generation of a very simple model: the majority of what a T cell “sees” are those amino acids in a peptide that protrude out of the MHC binding groove and can contact the TCR. This rule did not have to be true, as we also found evidence that the amino acids that anchor the peptide in the MHCII groove can lead to different conformations of the TCR contact amino acids, and some TCRs can detect these differences. However, the general model was very recently substantiated by the work of Garcia and colleagues, when they developed technology to screen an estimated 10\(^8\) p:MHCII complexes against TCRs with known p:MHCII ligands to determine breadth of cross-reactive of MHCII-bound peptides an individual TCR could see (Birnbaum, Mendoza et al. 2014). This group came up with a very similar rule: TCR contact amino acid identity with a known peptide ligand is the simplest way to predict if a second MHCII-bound peptide will be able to bind the same TCR. The fact that this result was achieved in an independent assay for two different MHCII allelomorphs suggests that this rule could be applied broadly.
We were specifically interested in determining why naïve CD4\(^+\) T cell populations specific for foreign peptides bound to the mouse MHCII molecule I-A\(^b\) varied over two orders of magnitude in size. We hypothesized that since none of the foreign peptides we tested shared all 9 core amino acids with mouse self peptides, negative selection on foreign p:MHCII-specific cell populations must be based upon TCR cross-reactivity with non-identical self peptides. Using the rule that amino acid identity at at least 4 out of 5 TCR contact amino acids at P2, 3, 5, 7, and 8 was necessary to predict TCR cross-reactivity, we found an inverse correlation between the naïve population size of 21 foreign p:MHCII-specific CD4\(^+\) T cell populations and the number of I-A\(^b\) binding homologous mouse peptides. These results shed some light onto the findings of Moon et al. (Moon, Chu et al. 2007), who developed the p:MHCII tetramer-based cell enrichment strategy to perform the first in-depth analysis of the pre-immune CD4\(^+\) T cell repertoire and its effects on immune responses.

Chapter 4 addressed a separate but related issue – how does a large CD4\(^+\) T cell population behave during infection (Pepper, Linehan et al. 2010). As the issue of CD4\(^+\) T cell stability had not been addressed for a persistent bacterial infection where CD4\(^+\) T cells are the protective cell type, we studied p:MHCII-specific response to persistent *Salmonella* infection. Our results demonstrated that the CD4\(^+\) T cell response to a *Salmonella* peptide:MHCII was numerically and functionally stable for as long as the bacteria persisted. This stability was associated with peptide:MHCII-driven proliferation by a small number of T cells in the secondary lymphoid organs harboring bacteria. This
work is significant because it could explain why current vaccines targeting *Salmonella* or *Mycobacterium tuberculosis* wane in efficacy after antigen is cleared.

In chapter 5, the impact of a large population of CD4$^+$ T cells specific for a single p:MHCI was assessed. Remarkably, one large foreign p:MHCI-specific T cell population played a large role in controlling ST infection, despite the likely presentation of many other ST-derived p:MHCI. Putting these findings together, one could envision a day where immunologists are successfully able to predict immunodominant epitopes based solely upon knowledge of a bacterial proteome and an individual’s HLA type, then develop a safe and efficacious subunit vaccine that produces long lasting CD4$^+$ T cell-mediated protection by long term delivery of antigen.
References


