Display of Ovalbumin Peptides by Murine Class II Major Histocompatibility
Complexes and the Engineering of Responding T Cell Receptors

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

Benjamin Michael Roy

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

Professor Jennifer A. Maynard, Adviser

December 2011

Acknowledgements

My thanks go out to everyone who helped me along the way. Principally, to Professor Maynard for mentoring me and giving me the freedom to pursue interesting questions. I would also like to thank the members of my committee for their time and efforts. All the current and former students with whom I have worked have deserve thanks for input regarding the research as well as setting up and maintaining the lab. Longtime colleague and friend Ryan Myhre, in particular, stands out in this regard.

Abstract

During an induced immune response, the BALB/c strain of laboratory mouse presents ovalbumin peptides by major histocompatibility complexes that stimulate responding T cells. This work first investigates the nature of ovalbumin peptide 323-339 presentation by the Class II MHC I-A^d in an *in vitro* setting, then focuses on the engineering of an individual T cell receptor (DO11) involved in ovalbumin response, and finally, pursues isolation of other responding T cell receptors with modern high throughput sequencing techniques.

T cell activation requires formation of a tri-molecular complex between a major histocompatibility complex (MHC), peptide, and T cell receptor. In a common model system, the ovalbumin epitope 323-339 binds the murine class II MHC, I-A^d in at least three distinct registers. The DO11 T cell recognizes the least stable of these, as determined by peptide-MHC dissociation rates. Using synthesized peptides in combination with IL-2 secretion assays, we show that the alternate registers do not competitively inhibit display of the active register four. In contrast, this weak register is stabilized by the presence of n-terminal flanking residues active in MHC binding. DO11 is sensitive to the presence of wild type residues extending to at least the P-3 peptide position, and the transfer of flanking residues in the P-4 and P-2 positions to an I-A^d presented hen egg lysozyme epitope increases the activity of that epitope substantially. Further, we provide evidence that display of the active register is dependent on the stability of an adjacent, overlapping register three. When epitopes were inserted in a maltose binding protein carrier, activity positively correlated with stability of the adjacent register, which was modulated by P1 anchor residue substitutions. Collectively, these results illustrate the potential of flanking residue modification to increase the activity of weakly bound epitopes. This represents an alternative to substitution of anchor residues within a weakly bound register, which we show can significantly decrease the activity of the epitope to a T cell responder.

In order to monitor the presentation of ovalbumin peptides, we have engineered a high affinity, multivalent single chain T cell receptor complex that exhibits specificity towards the OVA 323-339 peptide when displayed by I-A^d. DO11 was expressed as a single chain T cell receptor at levels in excess of 30 mg/L when oriented as beta variable region followed by alpha region separated by a flexible linker and incorporating previously identified solubility-inducing mutations. High affinity variants of DO11 were obtained from a randomized CDR3α library displayed as pIII fusions by M13 bacteriophage and panned against the DO11-specific, pMHC-mimicking antibody KJ1-26. ScTCR dimers were created by introduction of a hexapeptide EE-epitope tag into the flexible (G$_4$S)$_4$ linker, followed by complexation with a specific monoclonal antibody. Tetramers were formed by further association of dimers with a fluorescently labeled α-IgG1, and shown to specifically target peptide-loaded lymphoma cells in flow cytometry.

The applicability of scTCR protein expression and phage display techniques is also explored with other TCRs and found to be generally applicable to those containing mouse TRBV13-2, a gene previously noted for its preferential usage in the responding repertoire in mouse autoimmune models. We also show that 172, a TCR isolated from a mouse experimental autoimmune encephalomyelitis model, can be expressed at high levels as a fusion with the short peptide IT9302, a peptide from the c-terminus of interleukin 10 shown to possess many of the immunosuppressive characteristics of its parent cytokine.

We also report on the adaptation of phage display vector pMoPac24 to display fusion proteins with M13 major coat protein pVIII along with three previously identified display-optimized variants. We find that the production of phage particles containing these fusions and surface display is specific to the variant being expressed, but we demonstrate more efficient display for pVIII fusions as compared to pIII for DO11, as measured by inhibition of Interleukin 2 (IL-2) secretion during a 24-hour incubation with antigen presenting cells and

specific peptide. The use of these vectors may be necessary for directed evolution of T cell receptors exhibiting low affinity interactions with their target peptide/MHC complexes.

Finally, we report on the high throughput sequencing of mouse T cell receptor genes using pyrosequencing technologies. Alpha and Beta primer sets were designed to amplify cDNA from isolated mouse total RNA. The set was validated on T Cells isolated from mouse spleens, and used to identify T cell receptor genes from two clones, 3B6 and 4C4. A high throughput comparison of ovalbumin immunized mice and unimmunized controls is in progress. Identification of and subsequent engineering of T cell receptors involved in the ovalbumin response may provide further tools for future research on this murine system.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Ovalbumin has long been employed in studies of the murine immune system because of its widespread availability. This work examines a very specific slice of protein-protein interaction arising after ovalbumin immunization within a much larger series of immunologic events. First examined is the interaction between a small fragment of ovalbumin (residues 323-339) and the protein displaying that antigen, class II major histocompatibility complex (MHC) I-A^d on specialized antigen presenting cells (APCs). Next, the DO11 TCR responding to that specific pMHC is engineered for production in soluble form and for binding to its target with higher affinity. Finally, methods for high throughput sequencing of the ovalbumin-responsive TCR repertoire are reported, such that other TCR candidates for engineering might be identified, and that the data might give us greater understanding of the larger immune response. To place these pursuits in their proper context, a brief overview of the T cells function within the immune system will initially be presented, followed by an examination of selected developments in recombinant protein that are specifically applied here to TCRs.

1.1. T Cells in the Adaptive Immune System

T cells comprise one branch of the body's adaptive immune system. The immune system consists of "lymphatic" tissues and vessels that provide access for its components to monitor and combat disease. The innate immune response consists of a variety of cells and proteins adept at identifying pathogens through common motifs, mobilizing effectors, and stimulating adaptive immune responses.

The adaptive immune system is the body's specialized defense system against the rapidly evolving world of disease-causing agents. When properly stimulated, it can specifically recognize and eliminate a broad spectrum of viruses and cancerous cells. The system is replete with checks and balances that mediate its aggression. Improper stimulation can cause diseased cells to be ignored (e.g. latent tuberculosis or HIV) or healthy tissues to be attacked (autoimmunity). Much of the potency of the adaptive immune system lies in the range of T cells and B cells functions. Adaptive responses are antigen-specific: T cells and B cells possess tremendous diversity in their receptors, which allows them to carry out a tailored defense against viral or bacterial threats. In addition, T cells play integral roles in responses unrelated to specific pathogens, as in autoimmunity, cancer, and transplant rejection. T cells act through the recognition of peptides (protein fragments) displayed on cell surface major histocompatibility complexes. Based on the nature of TCR interactions with peptide/MHC (pMHC), and the environment in which it is received, the cell may receive a variety of distinct signals that in turn, coordinate other aspects of immune response (Figure 1.1).
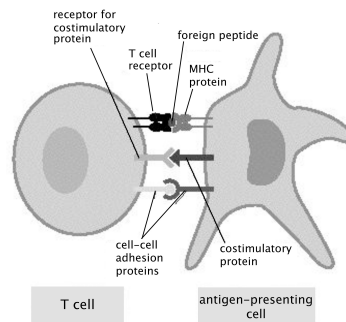
Figure 1.1. T Cell Interaction with an Antigen-Presenting Cell.

T Cells recognize peptides in complex with MHC on the surface of antigen-presenting cells. Activation is dependent on peptide recognition as well as the interaction of costimulatory receptors. Cellular interaction is stabilized by adhesion proteins. (Source: (1))

2

The majority of the T cell repertoire is composed of two different types: CD8[+], or cytotoxic T cells, and CD4[+], or helper T cells. CD8 and CD4 themselves are co-receptors, surface molecules that are thought to facilitate the interaction of T cell receptors by binding to constant regions of an MHC and stimulate intracellular signaling pathways. CD8[+] T cells specifically bind to MHC Class I, and CD4[+] T cells to MHC Class II. Activated CD8[+] cells generally respond with a cytotoxic barrage. They release cytokines and granzymes that destroy the target cell. CD4[+] cells, on the other hand, orchestrate an effector response that is carried out by other cells, such as macrophages and B cells. Within CD4[+] and CD8[+] T cells are subtypes with distinct behavioral patterns. More detail on these types and the criteria for activation of T cells is presented below. In order to understand the distinct functions of CD4[+] and CD8[+]T cells, the role of the MHC in antigen processing must be understood.
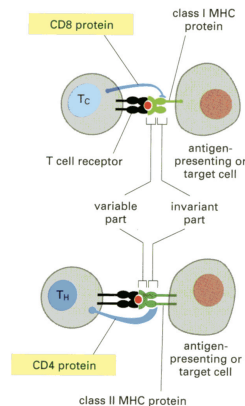
Figure 1.2. Coreceptors Facilitate T Cell Binding.

CD4 and CD8 Proteins interact with non-polymorphic regions of MHC during T cell antigen recognition. (Source: (2))

1.1.1. Antigen Presentation

Major histocompatibility complexes are glycoproteins classified as either class I or class II. MHC class I display peptides from within a given cell in a binding pocket composed of a single protein. Virtually all cells (including

3

professional antigen-presenting cells) express MHC I. As proteins are synthesized within cells a fraction are degraded; the resulting peptide fragments are complexed with MHC I molecules in the endoplasmic reticulum. These complexes migrate to the cell membrane, where they may interact with scanning CD8$^+$ T cells. MHC class I thus present a cross-section of the proteins produced inside a given cell, including foreign proteins present in an infected cell. Class II MHCs are heterodimers, whose binding pockets are made up of two distinct proteins. Unlike class I, class II MHCs are presented only by specialized antigen-presenting cells: macrophages, dendritic cells, and B cells. These cells internalize and degrade extracellular material, displaying the resulting peptide fragments and exhibiting any proteins that may be present in infected or inflamed tissue. MHC I and II also differ in the lengths of displayed peptides. The MHC I peptide binding pocket is closed, and displays peptides of only 8-10 amino acid residues in length. MHC II pockets are open ended, allowing long peptide fragments to overhang. There is theoretically no limit to the size of peptide that may be displayed by class II MHC. In practice, overly long peptides are trimmed, and the complex typically displays peptides made up of 13-17 amino acids (3).
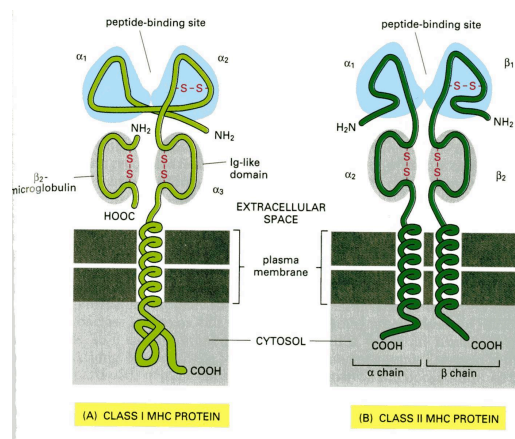
Figure 1.3. MHC Structure.

MHC Class I (*Left*) and Class II (*Right*) are membrane bound proteins that display peptide fragments in their outermost domains (α1 and α2 for MHC I, α1 and β1 for MHC II). (Source: (1))

The ability of an MHC to display a peptide "epitope" is dependent on non-covalent amino acid contacts between the peptide and its pocket-like binding site. MHCs are polygenic and polymorphic, meaning that several different genes code for MHCs, and many variations of these genes exist. All told, heterozygous humans can possess six different MHC alleles, (two each of HLA-DR, -DQ, and –DP) and mice four (two each of I-E and I-A). This genetic variety increases the landscape of peptides represented in MHC display (for review: (4)).

1.1.2. T Cell Development

T cells originate in the bone marrow as undifferentiated stem cells. Some of these cells migrate to the thymus, a lymphatic organ in the chest cavity, where they are differentiated toward a T cell lineage in the thymic stroma. In the thymus, the cells undergo genetic rearrangement to form functional T cell receptors in a process known as VDJ recombination. Each T cell receptor is a heterodimer consisting of an "α" and "β" chain (Figure1.4). Gene segments in each individual T cell rearrange to create a single gene for each chain, which, after transcription, combine to form a unique receptor (Figure 1.5). Each chain consists of a variable and constant (or germline) region. Variable regions are located at the interface of variable (V), diversity (D), and joining (J) gene segments for β chains (α chains contain no "D" segments). Additional diversity in TCRs is achieved through the incorporation of random nucleotides during gene rearrangement, known as p- and n-nucleotide addition. The total theoretical diversity of human T cells is estimated at approximately 10^{18} unique TCRs (3). The recognition of peptide/MHC takes place largely through three loops present on both chains that contact the binding partner. These loops are known as complementarity-determining regions (CDRs). More detail on antigen-recognition is discussed below.
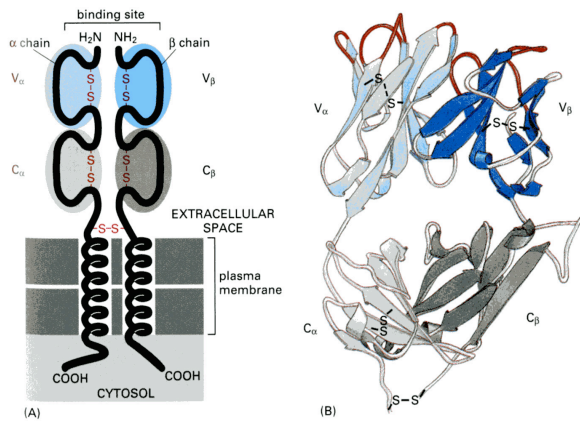
Figure 1.4. T Cell Receptor Structure.

Each TCR chain consists of a constant and variable region (*Left*). Disulfide bonds in the variable region expose 3 loops (*right*, in red), known as complementarity-determining regions, to peptide/MHC for recognition. (Source: (1))
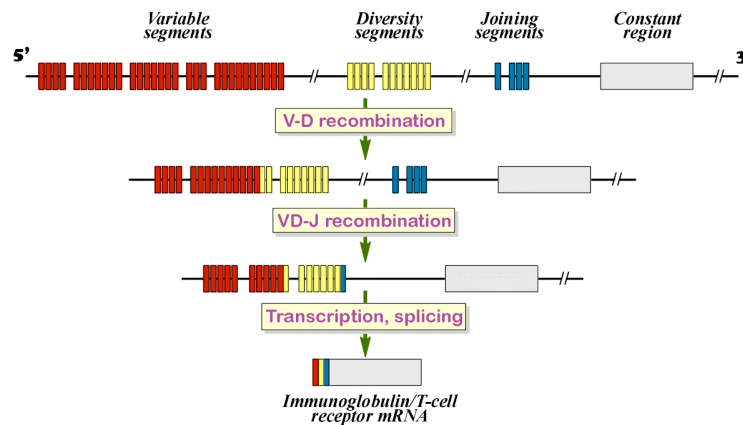
Figure 1.5. VDJ Recombination.

Diversity in T cell receptors is created through the random, but stepwise pairing of V-D-and J gene segments. (Source: (5))

T cells are functionally selected in the thymic stroma through the processes of positive and negative selection. In order survive, developing cells must recognize peptides in the thymus, but not interact too strongly. In the thymus, about fifty million thymocytes are generated each day, with only 2-4%

continuing on to the body as mature T cells. T cell progenitors in the thymus require stimulation to survive, but thymic stromal cells present only native, or "self" peptides to these thymocytes (Figure 1.6). Most developing T cells cannot recognize self-peptide/MHC, and thus do not receive the stimulation necessary for survival. These cells continue to rearrange their receptor genes until a "functional" TCR is produced, or die idly. This process, known as positive selection, orients the T cell repertoire to recognize an individual's personal set of MHC alleles. An estimated 10-30% of rearranged T cells bind successfully to self-peptide/MHC, and are positively selected (3).
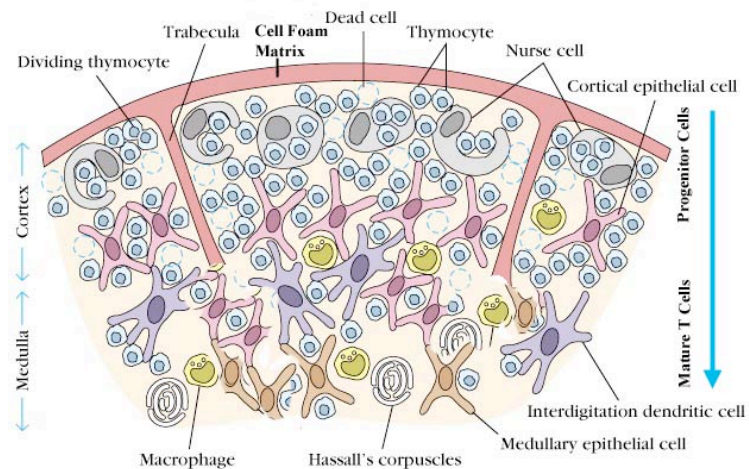
Figure 1.6. Developing T cells in the Thymus.

T cells encounter self-antigen presented by cortical epithelial cells and antigen presenting cells during maturation in the thymus. (Source: (6))

In order to survive, however, cells must also avoid negative selection, the process that removes self-reactive T cells. In addition to thymic stromal cells, developing T cells encounter antigen-presenting cells (APCs) that originate in the bone marrow. Apoptosis (programmed cell death) is induced in T cells that strongly interact with these cells. Strong self-reactivity in mature T cells could predispose an individual to an autoimmune response.

The mechanisms by which positive and negative selection occur are not completely delineated. The different cell responses, *i.e.* stimulation or apoptosis, are thought to be a function of signals propagated by the T cell receptor, and are almost certainly also a function of secondary signals in the thymus. In addition to positive and negative selection, some developing T cells will become regulatory T cells, which will be discussed below.

Thymic education of T cells ensures that the peripheral repertoire will largely be able to distinguish self from non-self, despite having been exposed essentially only to thymic tissue. Two distinct characteristics enable this efficient education. Typically, gene regulation is very tissue- and developmentally-specific. A given human cell expresses a limited number of proteins based on the function of its particular cell type. However, thymic stromal cells exhibit indiscriminate gene expression. A transcriptional regulator known as autoimmune regulator (AIRE) controls the expression of genes common to various cell types and ensures comprehensive expression of genes from these tissues. The other contributing factor to efficient thymic education is its architecture. Cells are closely grouped, forcing a developing cell to come into contact with many self-peptide/MHC. A developing cell spends an average of a week in the process of thymic education. The entirety of the process of weeding out self-reactive cells in the thymus is known as central tolerance. Despite its efficiency, self-reactive cells invariably remain, left to be controlled by peripheral tolerance mechanisms, as discussed below (7). First, consideration is given to how T cells are activated.

1.1.3. T Cell Activation

Following thymic development, mature T cells enter circulation. The "periphery" of the lymphatic system is defined as all lymph nodes and vessels outside of the thymus and bone marrow. T cells that mature in the thymus exit to the periphery, where they circulate via blood and lymphatic vessels between various lymph nodes. Peripheral cells require stimulation to survive, in a process

similar to that of positive selection in the thymus. In the absence of sustaining stimulation or activation by cognate antigen, the cell will die. Turnover of one to two million cells a day is common (3).

Activation of T cells occurs in lymphoid tissues and is dependent on the recognition of cognate antigen and costimulatory signals. Dendritic cells, B cells, and macrophages ingest extracellular antigen, process intracellular peptides and display them on surface MHC. Stimulated antigen presenting cells stop ingesting extracellular materials, start displaying large amounts of peptides, and migrate from inflamed sites to lymphoid tissues. Activated APCs also display B7 molecules, which are responsible for the costimulation of T cells. When naïve T cells recognize antigen via the TCR and B7 costimulatory signal via CD28, they begin to produce large amounts of growth factor IL-2, and divide rapidly into effector cells. In the absence of costimulation, as in signaling through unstimulated APCs or peripheral cells, T cells become anergic (discussed below). Activated T cell clones, in contrast, often no longer require costimulation to carry out effector functions (3).

Upon activation and clonal expansion, armed effector T cells respond based on their cell type. Activated CD4$^+$ T cells stimulate immune response through the release of cytokines, and the activation of effectors such as macrophages and B-cells. Upon stimulation, helper T cells differentiate into Th1 or Th2 cells, which shift immune response towards being cell-based, or humoral-based, respectively (Figure1.7). Differentiation is based on characteristics of the pathogen and the stimulating APC.
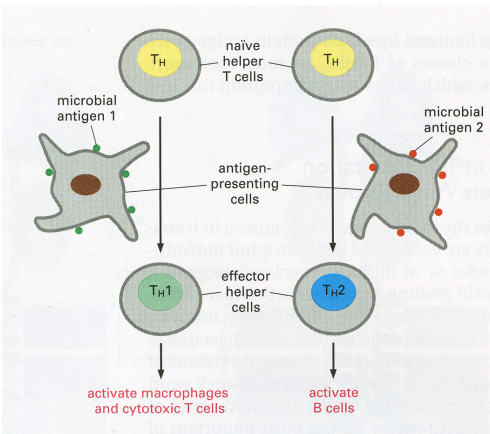
Figure 1.7. Stimulated CD4$^+$ T cells Adopt a Th1 or Th2 Type.

Upon encounter with antigen, antigen presenting cells produce a variety of cytokines that signal responding T cells to coordinate a Th1 (inflammatory) or Th2 (humoral response). (Sourece: (1))

An inflammatory environment and the presence of activated macrophages and IgG characterize cell-based immunity (Figure 1.8, left). Th1 cells release inflammatory cytokines such as IFN-gamma and TNF-α that drive this response. Humoral responses are characterized largely by the presence of antibodies, mainly IgM, IgA, and IgE, as well as anti-inflammatory cytokines secreted by Th2 such as IL-4, IL-10, and TGF-β (Figure 1.8, right). Inflammatory responses are typically associated with viral infections and autoimmunity, whereas humoral responses are associated with extracellular bacteria and allergy. The class of immune response generated can significantly affect the course of a particular disease. The impact of multiple sclerosis is thought to be mitigated by a shift from inflammatory to humoral response (8). Patient survival against mycobacterium leprae, on the other hand, correlates to an inflammatory Th1 response (9). In addition, environments are self-regulating, in that cytokines produced by Th1 suppress Th2 cells, and vice versa.

CD4 T cells:
peptide + MHC class II

| T_H1 cells | T_H2 cells |

cytokines — CD40 ligand — CD40 — intra-cellular bacteria — macrophage

cytokines — CD40 ligand — CD40 — bacterial toxin — antigen-specific B cell

Macrophage-activating effector molecules	Others	B-cell-activating effector molecules	Others
IFN-γ GM-CSF TNF-α CD40 ligand Fas ligand	IL-3 TNF-β (IL-2) CXCL2 (GROβ)	IL-4 IL-5 IL-15 CD40 ligand	IL-3 GM-CSF IL-10 TGF-β CCL11 (eotaxin) CCL17 (TARC)

Figure 1.8. Inflammatory vs. Humoral Immune Response.

Inflammatory immune response is driven primarily by Th1 activation of macrophages that have been infected or have ingested antigen. Humoral response is driven by Th2 activation of B Cells and is characterized by the secretion of cytokines that stimulate B cell growth. (Source: (3))

In contrast to the mostly indirect response of helper T cells, CD8[+] T cells are directly cytotoxic to infected cells in the periphery. In turn, CD8 cells have a higher threshold for activation. This threshold can be met through activation by mature dendritic cells, or through help granted by the presence of CD4[+] T cells. Cytotoxic effector molecules such as perforin and granzymes induce target cell lysis. CD8[+] T cells also display surface expressed Fas ligand, which can induce apoptosis in target cells.

1.1.4. Tolerance

Unrestricted activation of peripheral T cells would be destructive to the host, as self-reactive T cells invariably circulate. The process of thymic education

is not complete, and some self-reactive T cells escape deletion. In addition, many foreign antigens pose no threat, and the host is best served by ignoring them. Tolerance is the means by which the host avoids potentially harmful T cells responses; it can be induced in lymphoid or non-lymphoid tissues. Significant among the former are tissues that encounter a huge variety of foreign antigens as in the gut-associated lymphoid tissues. The mechanisms of tolerance induction are of significant interest in the study of autoimmunity.

Tolerance can be sub-classified into peripheral and central tolerance. Central tolerance refers to the thymus, and includes the previously described mechanism of thymic education as well as T regulatory cells that originate in the thymus. Peripheral tolerance refers to any tolerance independent of the thymus, including clonal deletion, anergy, and peripherally derived T regulatory cells.

1.1.4.1. Anergy and Clonal Deletion

Lymphocytes are said to be anergic when they become functionally inactivated after encountering an antigen, yet remain alive and resistant to activation. Clonal deletion, on the other hand, occurs when T cells die upon antigen encounter. Generally, naïve T cells that recognize antigens (foreign or self) in the absence of costimulation become anergic or are deleted. Necessary costimulatory signals are provided by the innate immune system, which depends on inflammation produced by motif recognition or tissue injury. The decision to become anergic or die probably depends on the nature of the TCR interaction as well as tissue-specific factors and levels of antigen-presentation (for review: (10)).

1.1.4.2. Regulatory T Cells

Regulatory T cells were initially thought to be a sub-population of anergic T cells that acquired the capability to induce anergy in surrounding cells (11). It has since become clear that regulatory T cells are neither strictly anergic, nor strictly generated in the periphery (12). However, the capability to subdue an

active immune response has made these cells an inordinately well-studied cell population and attractive target for immunotherapy. Various groups have presented evidence for seemingly distinct classes: Tr1 and Th3, which are generated in the periphery, and natural T regs, $CD4^+Foxp3^+$, which primarily originate in the thymus. The latter are the most well characterized group.

Natural T regulatory cells are helper (CD4) cells that have historically been marked by the presence of CD25. However, not all $CD25^+$ cells are regulatory, as this surface molecule is a receptor for growth factor IL-2 and is expressed on activated T cells. The most reliable indicator of this regulatory class of T cells is the expression of the Foxp3 transcription factor. This factor is thought to mediate a wide variety of cellular processes and is essential to the development of T regulatory cells.

Many, probably most, $CD4^+Foxp3^+$ T regs develop in the thymus as self-reactive T cells that have escaped deletion. Mice subject to a thymectomy prior to day 3 of life (when T regs first leave the thymus) develop autoimmune disorders (13). Natural T regs exert dominant immunosuppression in the presence of $CD25^-$ cells, which has made them an attractive population for immunotherapy (12, 14-16). The mechanism by which T reg cells act is as yet unclear, as divergent results have emerged *in vitro* and *in vivo*. In short, their mechanisms seem to be wide-ranging, and include direct interaction with responding T cells and indirect suppression mediated by interaction with antigen presenting cells. Immunosuppressive cytokines IL-10 and TGF-β play a role, as well as competition for binding with the IL-2 stimulatory cytokine, and expression of the CTLA-4 receptor, an inhibitory receptor for B7 costimulatory receptors. The development of T regs is influenced by high affinity TCR interactions with self-peptide (17-20).

1.1.5. Molecular Mechanisms of T cell activation

The interaction of T cell receptors with pMHC occurs through molecular

contacts in its six CDR loops (three per variable region). The TCR docks diagonally with the MHC such that the alpha chain loops are oriented towards the N-terminus of the peptide, and the beta chain towards the C-terminus. The peptide lies in a groove between parallel alpha helices of the MHC, and makes specific contacts with both CDR3 loops. Of the nine peptide residues within this binding pocket, typically at least three side chains are significantly solvent exposed, and are capable of forming specific van der Waals contacts and hydrogen bonds with TCR (for review, (21)).

TCR contacts with the MHC itself occur through interaction of CDRs 1 and 2; TCRβ regions interaction MHCα1 helix, and TCRα regions interact with the MHCβ2 helix (or MHCα2 in the case of class I complexes). Within this conserved arrangement, however, these germline-encoded TCR regions seem to possess remarkable polyspecificity (22). Variable gene usage does not seem to be restricted for particular MHC genes, and a single variable region may adopt different crystal contacts in its CDR1 and CDR2 depending upon the binding of the flexible CDR3 loop to the peptide.

This flexibility in recognition seems integral to the actions of T cells in the periphery. Recognition is hypothesized to occur by a two-step mechanism: CDRs 1 and 2 transiently binding the MHC and allowing the flexible CDR3s to alter conformation, forming specific peptide contacts (23). This mechanism would allow for T cells to "scan" presented peptides. There is evidence that continued recognition of the more prevalently presented self peptide-MHC complexes delivers a positive signal below the activation threshold that is required for survival of peripheral T cells.

Various models have been proposed to explain T cell activation based on a small number presented foreign peptides in the midst of many self-peptides. The serial triggering model suggests that each specific pMHC activates a number of specific TCRs (24). This predicts that an optimal strength of TCR-pMHC interaction exists, beyond which T cells are insufficiently activated. The kinetic

14

proofreading model emphasizes that a complex half-life of a certain duration is necessary for cooperative binding events necessary for signaling in T cell activation (25). Molecularly, there is considerable evidence that TCR signaling, mediated by the CD3 complex, involves the aggregation of various TCRs and their coreceptors and changes in localization of the cell surface surrounding the site of engagement (26).

1.2. Protein Engineering

Antibodies and T cell receptors (TCR) are the two principal targeting entities of the adaptive immune system, endowed with specific antigen recognition capabilities. Antibodies have found favor in the pharmaceutical industry, and along with that, extensive engineering platforms have been developed that have been drawn from in work with TCRs.

1.2.1. Monoclonal Antibodies

While the structural and genetic similarities of antibodies and T cell receptors are extensive, key differences include the ability of antibodies to recognize virtually any three-dimensional molecule (small molecule, lipid, carbohydrate, protein) with high affinity. TCRs, on the other hand, are exclusively membrane-bound, recognizing target peptides processed and presented by other cells with low affinity. Both possess substantial diversity due to gene rearrangement through VDJ recombination; receptors are formed through the recombination of disparate gene segments in the host cell's DNA. Unlike TCRs, antibodies undergo further rearrangement in the periphery in a process known as somatic hypermutation. By this process, individual nucleotides in the genes are randomly changed, followed by selection for high affinity variants (3). Antibodies are prominently used in diagnostics and therapeutics. As of 2008, 21 antibodies had been FDA approved as therapeutics, with over 400 others in clinical trials (27).

Antibodies are composed of two identical subunits assembled more or less in a "Y" conformation (Figure 1.9), with each branch capable of recognizing a target antigen. Each subunit is composed of a light and heavy chain, which are further subdivided into constant and variable regions. Variable regions are responsible for antigen recognition, while constant regions bind to other immune system effectors. As with the TCR, each variable region presents three solvent-exposed loops (complementarity-determining regions 1-3, CDR1-3) that contact the antigen, and define the specificity of the molecule.
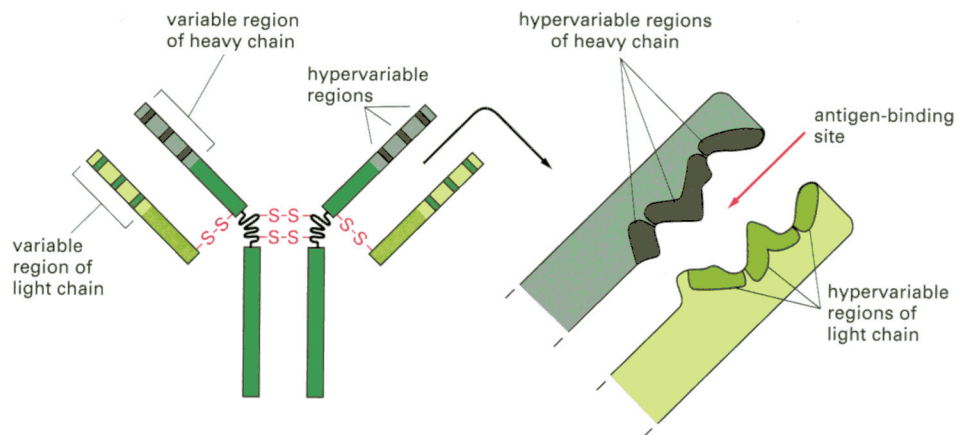
Figure 1.9. Antibody Structure.

(Source: (1))

Specific antibodies may be isolated in several ways. The classical method involves the creation of hybridomas. By this method, B cells are harvested from mice immunized with a particular antigen (28). These responding cells are fused with immortalized tumor cells; the resulting hybridomas retain the ability to produce antibodies as well as the immortal characteristics of the fusion partner. From a pool of hybridomas, specific cells are selected secreting genetically identical monoclonal antibodies (mAbs). These are likely to possess high affinity and specificity, as they have been subject to the physiological refining process of

16

somatic hypermutation. Because they are derived from a mouse, mAbs are highly immunogenic when used therapeutically in humans. Immunizing humans for antibodies is in many cases unethical and has the potential to induce harmful side effects, as in immunization with self-antigens. However, antibodies against self-antigens are often desirable, for instance, as targeting agents for cancerous cells. These concerns have led to the development of strategies to minimize immunogenicity while retaining the advantages of hybridoma technology, namely chimerization and humanization of murine antibodies.

Chimerization involves appending murine variable regions to human constant regions (29). Alternatively, antibodies may be humanized by grafting only the mouse CDRs onto a human antibody "scaffold" (30). To retain their potency, however, scaffolds generally must be modified so that displayed loops retain their native conformations.

As an alternative, murine involvement may be bypassed altogether through the use of naïve and synthetic libraries (31). Naïve libraries are harvested from the blood or other peripheral immune locations of humans who have not been immunized against a particular antigen. Because they lack somatic hypermutation, antibodies isolated from these libraries are typically of lower affinity.

Synthetic libraries take advantage of developments in scaffolding templates and phage display to provide large, specific libraries. Typically, synthetic libraries are created by fusing fully randomized complementary-determining loops to stable scaffolds. In the Human Combinatorial Antibody Library Project (HuCAL), a stable scaffold was creating using consensus framework regions, with randomized CDR3 introduced with oligonucleotides (32). Many antibody variants with high affinity and selectivity for chosen ligands have been selected from this library. Well-designed synthetic libraries can produce high affinity molecules, including antibodies against self-antigens, which are deleted in physiological settings.

Naïve and synthetic libraries are frequently combined with phage display technology to identify lead variants and then improve their affinity for ligand. This process is discussed further in section 1.2.3 and chapter 3.

1.2.2. TCR-like Antibodies

Limitations in the TCR repertoire have led to attempts to develop antibodies that recognize specific peptide/MHC complexes (for review (33)). Antibodies are naturally of higher affinity, in part because the previously discussed processes of thymic education that limits TCR affinity, as well as the lack of somatic hypermutation in TCRs. In addition, antibodies are naturally produced in a soluble, stable conformation, whereas the development of soluble versions of normally membrane-bound TCRs has been slower. However, the development of peptide-specific antibodies has proven difficult as well, as early reports indicated these were primarily reactive with the MHC (34-37).

Unlike antibodies, TCRs are genetically encoded to discriminate specifically for a specific peptide in the context of an MHC. Given the difficulty in isolating pMHC-specific Abs, it has been suggested that antibodies are, in contrast to TCRs, genetically predisposed to ignore peptide/MHC, as no peptide antigens are present to provoke negative selection physiologically, and recognition could result in an autoimmune reaction against antigen presenting cells (38). Indeed, TCR-like antibodies isolated from large naïve libraries have been reported as having greater peptide specificity (39-42). Even in the absence of strong genetic predisposition against selection, the small surface area of exposed peptide (~25% of total pMHC binding site) represents a daunting target for libraries of variants capable of recognizing any three-dimensional structure.

Two crystal structures for TCR-like antibodies have been reported. Hyb3, specific for a tumor-associated melanoma antigen, was shown to bind in a distinctly MHC-specific manner, with only four of its six CDRs active in binding

(43). In contrast, the 25-D1.16 antibody specific for an ova peptide bound to class I H-2K^b adopts a canonical binding mode seen in TCRs (44).

The future utility of TCR-like antibodies as compared to TCRs themselves will depend on the ability of each platform to overcome its own hurdles: TCRs in producing soluble, high affinity reagents, and TCR-like antibodies in reliably targeting specific peptides.

1.2.3. Directed Evolution

Directed evolution is the process by which a protein is modified for a desired characteristic. It involves the generation of a mutant library, packaging into a host for display, and selection based on biochemical or biophysical properties of interest. Reported applications include the mapping of epitopes involved the interaction of specific proteins (45-47), generation of novel protein function (48, 49), optimization of metabolic pathways through modification of promoter or transcription factors (50, 51) and enzyme engineering, such as modulation of substrate use (52) or generation of novel catalytic activity (53). The basis for any selection mechanism is the connection between DNA and protein, such that genetic variants responsible for phenotypic change may be predictably isolated. This is most commonly accomplished by fusing the protein of interest to a naturally displayed host protein. Selection is typically accomplished through several successive rounds of sorting and amplification (Figure 1.10).
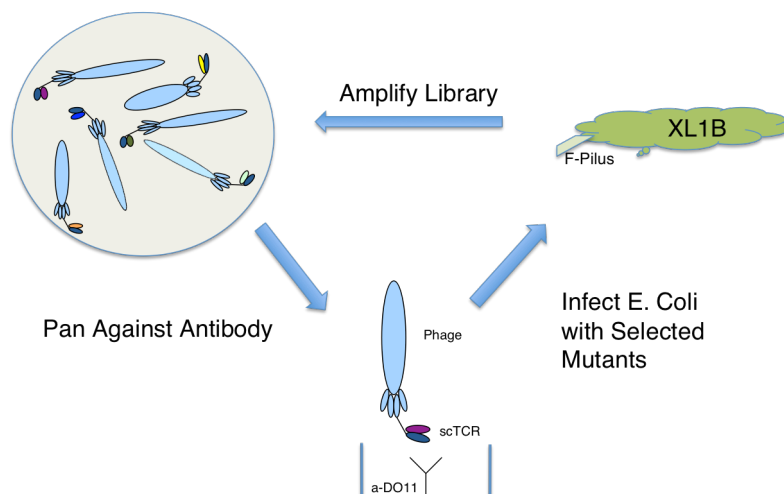
Figure 1.10. Phage Panning.

In this work, a T cell receptor is selected for increased affinity by means of bacteriophage display. Phage and their bacterial hosts are robust, with selection and subsequent recovery procedures simple and reliable. In addition, phage routinely encapsulate libraries in excess of 10^{10} unique variants (54). Various alternatives exist. Anchored periplasmic expression (APex (55)) also utilizes a bacterial host, but fused proteins are immobilized along the inner membrane, and require outer membrane disruption for exposure. Variants are selected by fluorescent sorting with a tagged ligand, and recovered via PCR amplification. In yeast display, eukaryotic yeast are used as an expression host, which facilitates the screening of more complex proteins requiring post-translational modifications (56). As with APex, variants are often sorted using flow cytometry, yet PCR is not necessary for identifying selected mutants, as intact yeast can be propagated post-selection. Typical library sizes associated with yeast display are between 10^5 and 10^7 mutants (57), as transformation of a library into yeast is generally less efficient than bacteria.

As an alternative to fusion proteins, ribosomal display takes advantage of the observation that in the absence of termination during translation, a free ribosome maintains the association of a nascent protein with its respective

mRNA (58). This process utilizes an *in vitro* transcription and translation system to avoid the use of a host, which often limits the diversity of the library (59). Only the number of ribosomes present restricts the diversity of ribosomal display. However, mRNA is very unstable compared to DNA, and nascent protein is necessarily connected with the bulky translation machinery during selection (60).

1.2.3.1. M13 Bacteriophage

Bacteriophages are viruses that infect bacterial cells. They are estimated to be the most abundant form of "life" on earth, with on the order of 10^8 found per milliliter of ocean water (61). One particular form, the M13 filamentous bacteriophage has been co-opted for the purposes of protein engineering (62). M13 is a non-lytic phage that has proven adept at displaying proteins ranging in size from small peptides to large immunoglobulins. Phage display is predicated on the physical connection genotype and phenotype in simple packaging. This allows for desirable DNA variants to be functionally selected and amplified. M13 particles are slender and long; they are made up of six different coat proteins that encapsulate single stranded DNA (Figure 1.11). The length of the particle ranges from approximately 700 nm to 2000 nm, with a diameter of only 6-7 nm. Protein VIII (pVIII) composes a majority of the particle, with thousands of copies encapsulating the drawn-out single stranded plasmid. At one end of the particle, approximately five copies of protein VI dangle off the end along with three to five copies of protein III (pIII), which mediates infectivity of the particle. At the other end five copies of proteins VII and IX reside, bound to the packaging signal on the encapsulated plasmid.
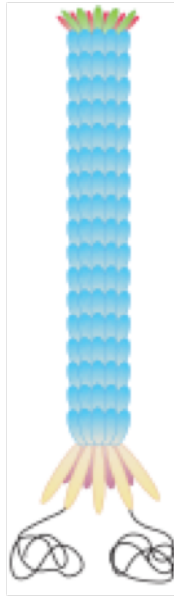
Figure 1.11. M13 Filamentous Bacteriophage.

Phage particles allow for elements of a DNA library to be functionally selected. DNA in the form of phagemid is stored within protein VIII coat; the gene coded for is displayed as a fusion with protein III. Other phage proteins may also be utilized in display. (Source: (63))

Phage particles infect bacteria via interaction with the bacterial F-pilus, a protein tube required for transfer of the F plasmid between adjacent cells. As the phage penetrates from the exterior of the cell through the bacterial and cytoplasmic membranes, its DNA is released from coat proteins, which migrate and become part of newly formed phage particles. The complement to the "plus-strand" phage-encapsulated DNA is then synthesized by bacterial replication machinery. This newly formed "negative" strand serves as the template for phage RNA production, which is then transcribed to produce viral proteins. Proteins II and X assist in replication of the viral DNA strand. Protein V interacts with the newly synthesized single stranded DNA, creating the drawn out structure that is incorporated in new phage particles.

Synthesized coat proteins migrate through the cytoplasmic membrane via the "sec" pathway to engage in particle assembly at the inner membrane.

22

Assembly is thought to be initiated by a "packaging signal" in the single-stranded DNA which binds to inner membrane-bound protein VII and protein IX. The DNA is then extruded through the membrane, with positively charged pVIII molecules interacting with the DNA to form the coat during transport. Assembly is completed when the DNA sequence ends, and proteins III and VI come into contact with this end of the DNA. In the absence of protein III to terminate assembly, the particle may take up another copy of the DNA and continue synthesis (64).

The M13KO7 phage used in this work contains a selectable kanamycin resistance gene as well as a modified packaging signal. Although it is capable of replication, a plasmid containing the wild-type origin of replication (phagemid) will be preferentially packaged into produced particles. Phagemid contain both phage and bacterial origins of replication, and can thus be subjected to modification by typical molecular biological techniques prior to their use in phage. The phagemid used in this work code for only a single bacterial protein. In order to package them into phage particles, helper M13KO7 phage must be added, which then synthesize the remainder of the proteins necessary for particle formation.

In this work, phagemid fuse proteins to copies of pIII or pVIII. For pIII, the gene is fused to a c-terminal fragment of the gene, such that it is incorporated into the surface of the phage particle in similar fashion to other copies of the protein III (65). This truncated version is devoid of infective properties, but due to its abundance, this is typically of little consequence. The presence of the protein is typically limited to a single copy among the three or four wild type pIII. In many cases, phage do not incorporate a pIII fusion at all, with some estimates placing this number at 90% or greater (66). Nonetheless, with M13 titers often exceeding 10^{11} per mL of bacterial culture, phage particles can easily accommodate large diversity libraries. pVIII display is discussed further in Chapter 3.

1.2.3.2. Library Creation

The introduction of genetic variation for selection can be carried out in random or site directed fashion, depending on the nature of the selection. For the alteration of well-defined binding sites, as is described in this work, a targeted approach is preferable. In the improvement of stability or expression levels, on the other hand, the targets are unpredictable, and a random approach may be more effective.

Random techniques include error-prone PCR and DNA shuffling. Error-prone PCR is a technique that modifies the standard polymerase chain reaction such that errors are purposefully introduced during the replication process (67). PCR depends on proper binding of complementary DNA sequences during thermocycling and the ability of the enzyme to properly replicate template DNA strands. The use of a polymerase that does not proofread, high repetitions of thermal cycles, or unequal ratios of the four component nucleotides ensure that a desired error rate is maintained in amplification. In addition, low annealing temperatures and high concentrations of magnesium ions facilitate annealing of DNA sequences that are not completely complementary. Error rate is dictated by the specified conditions; typically a mutation rate of between 0.1 to 2% of codons is desirable.

The use of error-prone PCR is limited to smaller genes. Multiple rounds of mutation and selection with error-prone PCR can lead to a buildup of total mutations, most of which are detrimental or non-functional. As these undesirable errors accumulate along with the beneficial ones, functional improvement of the protein becomes increasingly rare. For this reason, it is desirable to select only for functional mutants while weeding out neutral and negative mutants. Several techniques, most prominently DNA shuffling and StEP recombination, seek to create a combinatorial pool of variants and avoid buildup of deleterious mutations.

In DNA shuffling, a gene is broken up and recombined (68) through the use of DNaseI, a nuclease that digests DNA at random. The extent to which digestion occurs with this enzyme depends on the length of exposure. After a period of digestion, fragments are reassembled through a PCR reaction in which the fragments anneal to one another, and longer fragments that result are amplified by flanking primers. DNA shuffling is often performed starting from a pool of alleles, in attempts of preserving the useful mutants from each variant, though mutants generated from error-prone PCR may also be used. A single unmodified gene can also be used as a starting point, as the recombination process introduces a low level of point mutations (link).

Staggered extension process (StEP) recombination is an alternative to DNA shuffling that produces a combinatorial mixture without the use of a nuclease (69). The process involves replicating DNA as in PCR, except that abbreviated thermocycles are used such that a full-length sequence is achieved only after many cycles. During each cycle, fragments are extended slightly before being denatured and reannealed to a separate DNA template. A full-length gene is thus created from bits and pieces of multiple DNA templates. Unlike DNA shuffling, StEP requires an initial pool of variants to be used as template.

An alternative to the random error-based diversity creation is the introduction of targeted mutations with oligonucleotides (70). Short oligonucleotides bind to DNA and are incorporated in a newly created strand during PCR amplification. Oligos containing mutations sandwiched between strictly complementary flanking regions can introduce mutations into the target DNA. Moreover, a collection of oligos coding for all different amino acid sequences can introduce "random" mutations at a defined number of amino acid positions. Multiple oligos may be used to randomize disparate positions on a gene.

A single round of oligonucleotide amplification has the potential to completely randomize a series of amino acids. Randomized codons may be

indiscriminate (NNN) or designed to enrich for specific classes of amino acids based on patterns in the genetic code. A complete NNS codon, for example, codes for all twenty amino acids, but includes two fewer stop codons than a fully randomized NNN. Other codons exist for substituting groups of similar amino acids such as hydrophobic or hydrophilic residues (71).

Site-directed oligos can be used to amplify an entire plasmid linearly using a single oligo and single-strand DNA template, or exponentially, with double-stranded template and two opposite sense primers. Alternatively, oligos can be amplified using an overlap PCR strategy within a larger cassette that is later digested and ligated into the plasmid. Full plasmid amplification is preferable, as digestion and subsequent ligation are poorly efficient and can reduce the size of the resulting library. For this reason, even random methods have been adapted to full plasmid amplification, using cassettes up to 1 kb containing randomly distributed mutations as "mega-oligos" to amplify around the plasmid (Pai, in submission). Following full plasmid amplification, covalently closed-circular, double-stranded DNA (CCC-dsDNA) containing mismatched regions are be transformed into *E. coli*, which amplifies each strand individually. Pre-treatment strategies for the template DNA reduce the presence of the wild type after amplification. The template may be methylated, and treated with the CpG specific restriction enzyme DpnI after amplification. Alternatively, the template DNA can be grown in the presence of uridine in a strain deficient in thymine incorporation (*dut-,ung-*), and later transformed into a strain that preferentially inactivates uracil-containing dna (*ung+*).
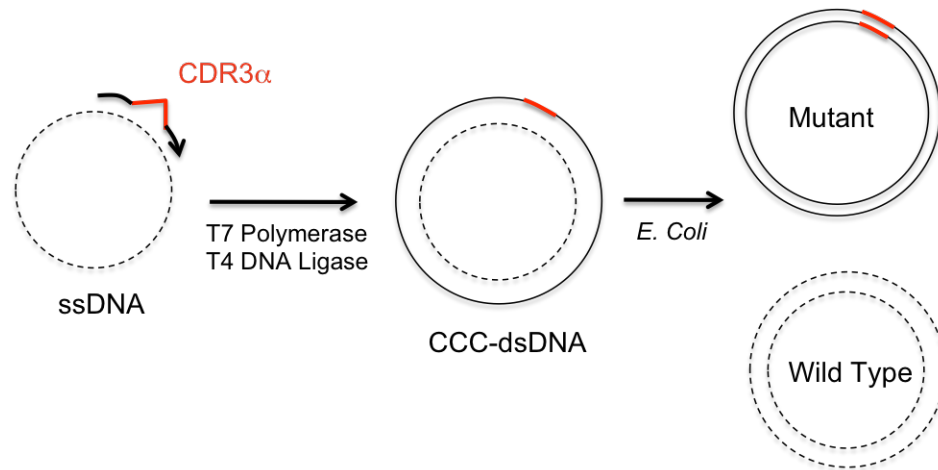
Figure 1.12. Oligonucleotide Based Mutagenesis.

Mutant oligos bind to DNA flanking CDR3α, the CDR itself containing randomized NNS codons. An enzymatic reaction creates a copy of the phagemid incorporating the oligo. Upon transformation into *E. coli*, the wild-type DNA with dUTP will be preferentially degraded; surviving colonies will comprise the library. (Adapted from: (54)).

In this work, library creation is performed according to the protocol of Sidhu and Weiss (Figure1.12, (72)). Oligos containing randomized NNS codons are used to linearly amplify single-stranded, uracil-containing template prior to transformation into high efficiency *E. coli*.

1.3. High Throughput Pyrosequencing

Traditional DNA sequencing (the Sanger method) is carried out by random incorporation of fluorescent-labeled, chain-terminating dideoxynucleotides during extension of an oligonucleotide via polymerase (73). DNA fragments are separated by length via gel electrophoresis, with the identity of each sequential nucleotide determined by the observed fluorophore signal. This technique requires a significant quantity of each gene to be sequenced, typically after incorporation into a plasmid.

In pyrosequencing, nucleotide addition to an extending chain is detected in real-time as the release of a pyrophosphate (74). Nucleotides of known identity are added to a reaction mix, and when incorporated, they release a pyrophosphate, which is subsequently converted to ATP by ATP sulfurylase. The ATP then provides energy for the oxidation of luciferin by luciferase, which releases photons in proportion to the number of nucleotides incorporated. In the solid phase reaction, excess nucleotides are washed away, while in the liquid phase version (75), excess nucleotides are degraded by apyrase.

A high throughput version of the pyrosequencing reaction has been commercialized by 454 Life Sciences (76). In this version, single stranded DNA is first labeled with a short identifying primer, and ligated to small beads. Beads are trapped in an oil-water emulsion in the presence of the required enzyme, and the reaction proceeds. The emission of light is monitored by a massively parallel fiber-optic system. From this data, nucleotide sequence can be determined. In the version of the instrument employed here, each individual reaction produces read lengths of up to 500 base pairs. Entire TCR variable region sequences (~300-350 bp) can thus be determined from a single reaction.

CHAPTER 2

THE ROLE OF FLANKING RESIDUES IN I-A^D ASSOCIATED

OVALBUMIN PRESENTATION TO DO11

2.1. Introduction

The adaptive immune system responds to foreign proteins through the use of cytotoxic (CD8$^+$) and helper (CD4$^+$) T cells that activate and direct other effectors to mount a protective response. These T cells are activated after a productive interaction between the membrane bound T cell receptor (TCR) and the surface of cells presenting a composite peptide/major histocompatibility complex (pMHC). Crystal structures have revealed regular patterns of peptide-MHC interactions [77], enabling prediction of peptide epitopes present within larger proteins. CD8$^+$ T cells recognize Class I MHC molecules displaying 8-10 residue long peptides derived from intracellular proteins, proteolytically processed to a defined length in order to fit within the closed MHC binding groove. In contrast, CD4$^+$ T cells recognize extracellular material that has been internalized, processed, and displayed as variable length peptides in complex with heterodimeric Class II MHC molecules on professional antigen presenting cells. Because of an open-ended peptide-binding groove, class II MHC typically display peptides with 13-17 or more residues, with the termini extending out from the binding groove [78]. These longer peptides present distinct challenges for epitope prediction since there are multiple potential modes of interaction with the MHC [79, 80].

Class II peptides bind the MHC in an extended, polyproline II-like conformation, flanked by two antiparallel α-helices, one from each MHC chain [81]. These helices form a conserved network of hydrogen bonds with the

peptide backbone spanning the length of the binding pocket that dictates a sequence independent conformation of the displayed peptide. The stability of a specific peptide is determined in part by side chain interactions with these helices as well as internal binding pockets. The peptide rests atop a series of pleated beta sheets, which form four prominent binding pockets, named for the respective position within the peptide of the residues that occupy them: P1, P4, P6 and P9. The stretch of nine amino acid residues spanning these pockets is referred to as the peptide core, and the residues that fill these pockets as anchor residues. Considerable effort in class II epitope prediction has focused on defining anchor residue preferences for specific MHC alleles (82). The remaining residues within the core orient their side chains toward the helices or away from the MHC, in order to interact with the TCR.

Within the longer peptides presented by class II MHC, multiple core sequences may overlap, such that a single pMHC complex may heterogeneously present multiple cores for display to responding T cells. Murine class II I-A alleles appear particularly susceptible to this phenomenon as anchor preferences are ill-defined among the diverse set of known binding peptides. The apparent degeneracy of I-A alleles emphasizes the need for epitope prediction models which incorporate the contributions of separate registers within an epitope. Indeed, early sequence-based prediction models found that incorporation of flanking residues outside of the predicted core improved predictions for HLA-DQ, human homologs of I-A (83). This appears partially attributable to flanking residues either interact directly with MHC or influence epitope processing (84). More recent models have continued to incorporate flanking residues in sequence based prediction models (79, 85, 86), with some explicitly including flanking residues for their roles in alternate register binding (79, 80).

The I-A allele is associated with numerous autoimmune mouse models, while the homologous HLA-DQ allele is associated with insulin dependent diabetes milletus, celiac disease, and potentially multiple sclerosis in humans

(87, 88). This allele-specific autoimmune susceptibility may result from peptide presentation in multiple registers (89). In this proposed mechanism, auto-reactive T cells escape thymic deletion because their target self-epitopes are hidden by "register masking," whereby they are presented less frequently than the more thermodynamically stable alternate registers within the same proteolytic fragment. Thus, T cells responding to the more stable registers are deleted, while those responding to less stable registers migrate to the periphery. At the site of inflammation, where the peptide is highly expressed, the weaker register will be presented with sufficient frequency to activate self-reactive T cells and induce disease. Specific epitopes identified in the NOD mouse model of type 1 diabetes (90) as well as experimental autoimmune encephalomyelitis (91-93) have been identified as potentially susceptible to this mechanism.

The DO11 T cell hybridoma is a popular model system used to study CD4$^+$ T cell responses, originally elicited by vaccination of BALB/c mice (H-2^d haplotype) with hen egg white ovalbumin (OVA). The specific peptide recognized was originally described as the tryptic fragment corresponding to residues 323-339 as presented by I-A^d (94). Subsequent studies first identified two distinct binding registers within this long peptide (324-332; 327-335) (95). Later, a third register was identified (329-337) and a fourth proposed (326-334; Table 2.1) (96). Among these, the DO11 T cell hybridoma was shown to recognize the *c*-terminal register (residues 329-337, referred to hereafter as register four) through the use of *N*- and *c* -terminal truncations and targeted amino acid substitutions (96). The failure of MHC tetramers with a covalently linked 323-339 peptide to stain DO11 and other peptide-specific T cells has driven interest in further defining the register recognized. Recently, Landais *et al.* synthesized four tetramers, each presenting one of the four respective nine residue core sequences present within the 323-339 peptide, and confirmed that only the register four peptide stained DO11 cells and bound the free T cell receptor (97).

Register			P-4	P-3	P-2	P-1	P1	P2	P3	P4	5P	P6	P7	P8	P9					
1 (324-332)						I	S	Q	A	V	H	A	A	H	A	E	I	N	E	A
2 (326-334)				I	S	Q	A	V	H	A	A	H	A	E	I	N	E	A	G	R
3 (327-335)			I	S	Q	A	V	H	A	A	H	A	E	I	N	E	A			
4 (329-337)	I	S	Q	A	V	H	A	A	H	A	E	I	N	E	A	G	R			

Table 2.1. Alternate Registers in OVA 323-339

Peptide-MHC complex stability correlates broadly with immunogenicity (98, 99), as stable epitopes are more prevalent on the cell surface, and subsequently more likely to be recognized by responding T cells. Stability is commonly approximated as the dissociation rate of fluorescently labeled peptide from soluble MHC, although this approach can be complicated by highly overlapping registers. McFarland and Beeson showed that OVA registers one and three are long-lived and stable (95). Supporting the thermodynamic stability of register one, a complex between the 323-339 peptide and I-A^d was crystallized in that register (100), and tetramers with long, unhindered covalent linkers primarily inhabit register one (97). In contrast, the DO11-responsive register four peptide has a very rapid off-rate, with a half-life of around twenty minutes (99). To minimize alternate registers confounding this measurement, the amino acid sequence flanking the register four core was truncated at P-2 on the *N*-terminus, with the P-2 position altered from valine to lysine, which is presumably too large to stably occupy a register one anchor position. These peptide dissociation rates predict that register four would behave as a masked epitope, albeit in the context of a foreign antigen. Paradoxically, Landais *et al*. found that pMHC tetramers presenting only the fourth register stained a significantly greater number of responsive T cells following OVA immunization than tetramers presenting the more stable first and third registers.

Here, we sought to quantify the extent to which registers one and three mask display of the active register four, using synthetic OVA peptides

incorporating truncations and/or mutations to alter presentation of a given register and, ultimately, enhance register four display. In addition, recombinant proteins harboring internal peptide variants were employed to capture effects associated with epitope processing. After incubation of DO11 hybridomas with corresponding B-cells and one of the antigen variants, we measured T cell activation in terms of IL-2 secretion, as the most sensitive determinant of register four's accessibility to responsive T cells.

We observed that presentation of register one, but not register three, is independent of register four activity. Disruption of register one through *N*-terminal peptide truncation had no impact on activity. Attempts to stabilize register three through anchor substitutions in the context of a carrier protein actually *increased* the activity of register four, suggesting that these residues may cooperatively enhance display, although the equivalent substitutions in synthetic peptides showed no clear effect. Moreover, we provide evidence that register four stability has been underestimated by traditional techniques that truncate flanking residues. The transfer of flanking residues P-4 through P-2 to an unrelated hen egg white lysozyme epitope, HEL11-25, increased the activity of that epitope dramatically, suggesting these residues independently stabilize peptide-I-A^d interactions. Conversely, additional attempts to stabilize ovalbumin by amino acid substitution within the core consistently decreased activity. Collectively, our results suggest that in the case of 323-339, flanking residues are vital to activity in a previously unforeseen fashion. These results have broad implications for class II epitope prediction algorithms and protein engineering efforts to modulate protein immunogenicity in therapeutics and vaccines.

2.2. Materials and Methods

2.2.1. OVA and HEL peptides

Synthetic ovalbumin 323-339 peptide (p323) was purchased from Anaspec (#62571, Fremont, CA). Custom ovalbumin (OVA) and hen egg lysozyme (HEL) peptide variants were synthesized by Peptide 2.0 (Table 2.2, Chantilly, VA). All ovalbumin variants were amidated at the c-terminus. All peptides were 95% pure by HPLC analysis or greater.

2.2.2. Chimeric malE proteins containing OVA epitopes

Fragments of ovalbumin were introduced into the maltose binding protein (*malE*) in a manner similar to that first described by Martineau *et al.* (101). These short peptide sequences were introduced via overlap PCR to replace *malE* amino acid residues 133-140, flanked at the n-terminus by the amino acids PDPGSGSG, and at the c-terminus by PDPGS. This extended linker confers higher protein expression, along with sufficient length at the n-terminus to preclude contributions from residues other than glycine or serine during MHC loading. The resulting chimeric proteins were named based on the numerical OVA sequence of the inserted peptide. For instance, a chimera containing all registers of the OVA peptide is termed *malE*_323-337. Non-wild type residues introduced into *malE*_327-337 are named with respect to register four positions, such that a residue change from valine to methionine at position P-2 is termed 327 V(-2)M. The modified *malE* genes were flanked with directional *Sfil* restriction sites and subcloned into pAK400 for periplasmic expression in *E. coli* (102). The plasmid was sequenced to confirm epitope presence, transformed into strain BL21, grown overnight at 37 °C in 1-4 L shake flasks. After replacing spent media with fresh terrific broth, protein expression was induced with 1 mM IPTG for 4-24 hours at 25 C. The cells were harvested by centrifugation and cell pellets were resuspended using ice cold 0.1 M Tris pH 7.4-buffered 0.75 M sucrose at 3

mL per gram of cell pellet. An osmotic shock to release the periplasmic contents was performed by addition or 3 mL 1.0 mM EDTA per gram of cell pellet and approximately 0.3 mg/mL lysozyme followed by the addition of 6 mM $MgCl_2$ on ice. After an additional centrifugation step, the supernatant was dialyzed against 10 mM Tris pH 7.4, 0.5 M NaCl with 12-14 kDa MWCO dialysis tubing. Chimeric proteins were recovered via immobilized metal affinity chromatography (IMAC) with the *c* -terminal His tag and size exclusion chromatography (Superdex 75; GE Healthcare) in HBS and stored at -80 °C prior to use.

2.2.3. Antigen presenting and T cell hybridoma cell lines

The murine A20 B cell lymphoma line presenting I-A^d MHC was acquired from ATCC (TIB-208). An A20 variant line with a mutated, non-functional HLA-DMα gene, 3A5 (103, 104) was transferred from Dr. Kenneth Rock, Dana Farber Cancer Institute, Boston, Massachusetts. The murine DO11 T cell hybridoma line was transferred from Dr. Philippa Marrack, National Jewish Health Center, Denver, Co. The murine HEL11-25 responsive T cell hybridoma (105) was a generous gift from Dr. Andrea Sant, University of Rochester Medical Center, NY. All cells were maintained in DMEM supplemented with 10% FBS, 50 Units penicillin/ 50 mg streptomycin per mL and 2 mM L-glutamine.

2.2.4. T cell Activation

Each well of a 96-well plate was seeded with 15,000 A20 cells, 15,000 T cell hydridoma cells and peptide, *malE* chimera protein or control peptide/ protein per well in DMEM media without L-glutamine, except where noted. After a 24-hour incubation at 37 °C, 5% CO_2, supernatant was transferred to a new plate and stored at -80 °C prior to IL-2 quantification assays. All peptide/ protein doses were performed in triplicate wells and all experiments replicated at least twice.

For *malE* chimeras and control protein experiments, the relevant protein was thawed, filter sterilized, and concentrated using 30K molecular weight cut-off centrifugal filter units (Millipore UFC903024). Proteins were then quantified via

absorbance at 280 nm using a predicted extinction coefficient and molecular weight, diluted to equal concentrations, the concentration confirmed and serially diluted as indicated in the figures. In each well, the volume fraction of protein in HBS solution added did not exceed 10%, and an FBS concentration of 10% was maintained in all wells. Proteins were pre-incubated for four hours with antigen presenting cells (A20 or 3A5) cells prior to a 24-hour incubation with T cell hybridomas. The IL-2 concentration in the supernatant was quantified using a matched pair ELISA quantified with an IL-2 standard as described (BD #555148). All experiments were performed at least twice, and all wells were measured in triplicate. Data was analyzed with Graphpad Prism 5 (La Jolla, CA).

Glutamine was excluded from experiments (except where noted) to minimize cell growth, as preferential growth was observed in inverse proportion to T cell stimulation. Since overall IL-2 release is affected by the ratio of T cells to B cells, and assays were performed only at a limited number of discrete doses, we reasoned that differential growth could potentially mask small differences in the activity of specific peptides. The absence of glutamine lowered overall IL-2 release and required larger concentrations of peptide or protein to reach a given release.

2.2.5. Predicted peptide-MHC docking

The coordinates for I-A^d and the OVA 323-334 peptide were extracted from PDB 1IAO. The ovalbumin peptide was extended from the c-terminus to residue 338 and manually extended such that backbone dihedral angles (ϕ, ψ) approximate those observed in an extended polyproline type II helix (-75, 150). Computational peptide docking was performed with ClusPro (106).

2.3. Results

2.3.1. DO11 Activation is Modulated by Epitope Anchor Residue Substitutions

The rapid dissociation of register four relative to the other registers within 323-339 led us attempt to increase activity through a corresponding increase in its affinity for I-A^d. This poor stability is likely due to the weakly binding anchor residues at ovalbumin positions A329, A332, I334, A337 (Table 2.1). Anchor substitution is a common technique for increasing the affinity of a given register; previously, Chaves *et al.* reported the relative stability of commonly observed anchor residues with I-A^d (107). In this and other work, the interactions of each anchor residues with its respective binding pocket were found to be independent of the overall sequence (80).

Two previous reports described anchor substitutions to stabilize register four. Robertson *et al.* replaced all four positions with the corresponding anchor positions from a sperm whale myoglobin epitope with high affinity for I-A^d, using the following substitutions: A329E, A332I, I334V, A337S (here referred to as peptide 323-339 EIVS) (96), and reported increased activity for the variant. Similarly, Lazarski *et al.* introduced the V327K substitution to prevent register shifting as lysine is too large to fit the P1 pocket, as well as A332I and I334V (327-339 KIV) (99) and reported recognition by the hybridoma. To assess the relative success of these attempts, we synthesized both variants, along with an *N*-terminally truncated version of the Robertson peptide (328-339 EIVS; Table 2.2) to remove any influence of these flanking residues. To assess peptide activity, DO11 hybridomas were incubated for 24 hours with I-A^{d+} A20 lymphoma cells in the presence of peptide, with the supernatant subsequently assessed for IL-2 presence by a matched pair ELISA as a measure of T cell activation.

Peptide			P-4	P-3	P-2	P-1	P1	P2	P3	P4	5P	P6	P7	P8	P9				EC50 (mM peptide)	95% Confidence	Max response (ng/ml IL-2)	95% Confidence
OVA #	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339					
323-339	I	S	Q	A	V	H	A	A	H	A	E	I	N	E	A	G	R	1.6	1.3-2.0	1.1	1.0-1.2
325-337			Q	A	V	H	A	A	H	A	E	I	N	E	A			1.7	1.2-2.3	1.1	0.9-1.2
327-337					V	H	A	A	H	A	E	I	N	E	A			12	7-23	0.7	0.5-0.9
328-337						H	A	A	H	A	E	I	N	E	A			ND		ND	
GGV328-337			G	G	V	H	A	A	H	A	E	I	N	E	A			6	4-10	1.1	0.9-1.4
GGS328-337			G	G	S	H	A	A	H	A	E	I	N	E	A			12	7-20	1.0	0.7-1.3
GGM328-337			G	G	M	H	A	A	H	A	E	I	N	E	A			6	4-9	1.0	0.8-1.2
327-339 KIV					K	H	A	A	H	I	E	V	N	E	A	G	R	ND		ND	
323-339 EIVS	I	S	Q	A	V	H	E	A	H	I	E	V	N	E	S	G	R	ND		ND	
328-339 EIVS						H	E	A	H	I	E	V	N	E	S	G	R	ND		ND	
MalE chimera	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339					
323-337	I	S	Q	A	V	H	A	A	H	A	E	I	N	E	A			13*	8-20	0.8*	0.6-1.0
325-327			Q	A	V	H	A	A	H	A	E	I	N	E	A			18*	12-25	1.1*	0.9-1.4
328-337						H	A	A	H	A	E	I	N	E	A			73*	26-201	0.3*	0.1-0.5
327-337					V	H	A	A	H	A	E	I	N	E	A			41	27-61	0.6	0.5-0.8
327 V(-2)M					M	H	A	A	H	A	E	I	N	E	A			18*	13-25	1.0*	0.8-1.2
327 A(1)M					V	H	M	A	H	A	E	I	N	E	A			36	14-93	0.6	0.3-0.9
327 V(-2)E					E	H	A	A	H	A	E	I	N	E	A			25*	15-44	0.7*	0.5-1.0
327 V(-2)M, A(2)I					M	H	A	I	H	A	E	I	N	E	A			ND		ND*	
HEL #				11	12	13	14	15	16	17	18	19	20	21	22	23	24	25				
HEL11-25				A	M	K	R	H	G	L	D	N	Y	R	G	Y	S	L	6	2-20	0.7	0.3-1.1
G11-25			G	A	M	K	R	H	G	L	D	N	Y	R	G	Y	S	L	7	2-23	0.7	0.3-1.1
QAV13-25			Q	A	V	K	R	H	G	L	D	N	Y	R	G	Y	S	L	3*	1-6	1.5*	0.9-2.1
GGS13-25			G	G	S	K	R	H	G	L	D	N	Y	R	G	Y	S	L	6*	2-26	0.5*	0.2-0.8

Table 2.2. Modified Ovalbumin and HEL Peptides And Associated EC$_{50}$ Values.

EC$_{50}$ and maximum response values from 3-parameter curve fits of aggregate data from multiple (at least two) experiments. Positional numbering is in relation to the active register (register four of OVA in this case).

38

Surprisingly, all three peptide variants drastically reduced T cell stimulation versus 323-339 (Figure 2.1A). Notably, the full-length peptide with anchor substitutions (323-339 EIVS) stimulated more effectively than its shorter counterpart (328-339 EIVS). Given the documented affinity increase of 327-339 KIV for I-A^d (99), the decrease in T cell stimulation with these variants likely results from interference in TCR/pMHC recognition due to minor structural variations.
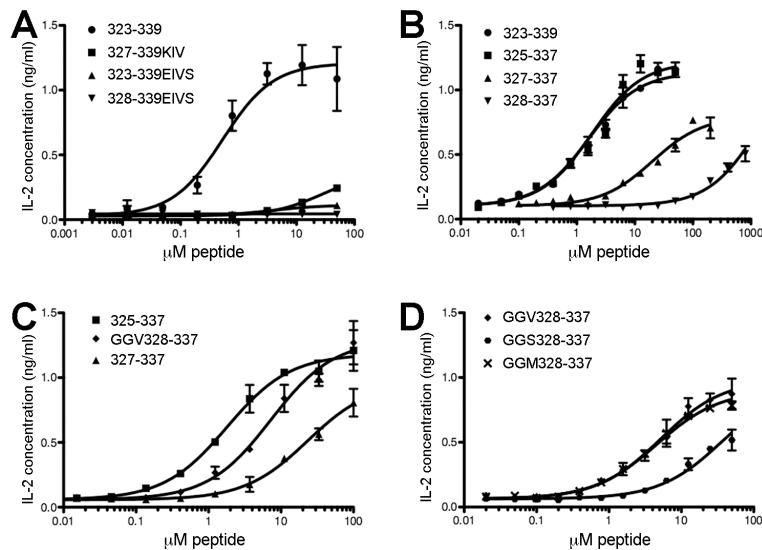
Figure 2.1. Specific flanking residues but not alternate register stabilization enhances OVA peptide activity.

DO11 hybridoma and antigen presenting cells were incubated overnight in the presence of exogenous peptide variants, followed by quantification of IL-2 released as a measure of T cell activation. **A,** Wild-type ovalbumin 323-339 (●) was compared with peptide variants modified at anchor residues P4 (A to I) and P6 (I to V) in conjunction with additional changes in an effort to stabilize the active register four. Peptide variants include 327-339KIV (■) with an additional *N*-terminal truncation and P-2 modification (V to K) to

39

prevent register shifting; 323-339EIVS (▲) the full-length peptide with all four anchor residues altered to strongly binding residues; and 328-339EIVS, an *N*-terminal truncation of the previous variant (▼). The cell culture media contained 2mM glutamine for this assay. *B*, In a subsequent effort to stabilize register four by removing alternate register anchors, the wild-type 323-339 peptide (●) was compared with *N*-terminally truncated variants 325-337 (n), 327-337 (▲), and 328-337 (▼). *C*, To determine whether peptide length is responsible for the activity differences observed in the truncated variants, the 325-337 variant (■) which includes register four and the *N*-terminal P-1 to P-4 residues was compared with GGV328-337, in which the P-3 and P-4 positions were replaced with glycine (◆) and a further truncated variant, 327-337 (▲), lacking the P-3 and P-4 residues. *D*, To determine whether binding of register three enhances register four activity, we changed the register three P1 anchor residue (also register four P-2) to amino acids predicted to stabilize (M) or destabilize (S) register three. We compared three variants with identical length: variant GGV328-337 (◆) with a wild-type V versus GGS328-337 (●) and GGM328-337 (✕).

2.3.2. DO11 Stimulation Modulated by Residues N-Terminal to P-2

The failure of anchor four residue substitutions to increase activity led us to explore *N*-terminal truncations as a means to destabilize alternate registers, which may compete with register four for MHC binding. The 323-339 peptide was progressively truncated from the n-terminus, with each truncated peptide extended to an amidated residue 337 and each peptide variant incubated with DO11 and A20 cells, as above. The 325-337 peptide, lacking the P-1 and P1 positions of register one, conferred identical stimulation in comparison to the 323-339 peptide, in terms of the 50% effective concentration (EC_{50}) and maximal response, suggesting that register one does not effectively compete with register four for display (Figure 2.1b, Table 2.2). Additional truncations to remove the residues at register four positions P-3 and P-4 reduce activity dramatically; the 327-337 peptide increases the EC_{50} nearly 10-fold, from 1.6 to 12 μM peptide while the maximal response decreases from 1.1 to 0.7 ng/mL IL-2. For the 328-337 peptide, activity is so severely reduced that accurate EC_{50} values could be

determined (Table 2.2). These data agree qualitatively with Robertson *et al.* who measured T cell proliferation from transgenic DO11 mice in response to a similar set of truncated peptides (96). In contrast to the theorized effects of alternate registers in autoimmune peptides, DO11 stimulation is maximal when the entire third register and at least one additional *N*-terminal residue are present, in addition to the active register four, which led us to explore the mechanism responsible for this enhanced activity.

2.3.3. Flanking Residue Transfer to HEL1125 peptide

A survey of the available crystal structures reveals various roles for *N*-terminal flanking residues. The T cell receptor, by virtue of its CDR1α and CDR3α loops, often contacts the P-1 position, and can interact with more distant residues (77). MHC contacts, on the other hand, appear to be more predictable, typically limited to canonical hydrogen bonds between residue P-1 and the side chain of β81H or Y, as well as residue P-2 and the main chain of α53 (108).

To evaluate the possibility that distant flanking residues of the ovalbumin peptide (325-337) were involved in T cell receptor contact, we employed a second model system, in which hen egg white lysozyme residues 11 through 25 (HEL11-25) are presented by I-A^d. This peptide has not been shown to occupy multiple frames in I-A^d, but has been characterized as a weak MHC binder with a half-life of approximately six hours (99), presumably due to a bulky arginine residue which is poorly accommodated by the small P1 pocket. We compared the activities of wild type HEL11-25 with variants in which the native P-4 to P-2 sequence was replaced with OVA-specific residues (QAV13-25) and a control in which these three residues were adjusted to GGS (GGS13-25). Since these changes lengthen the HEL11-25 peptide by a single amino acid, a control in which the HEL11-25 peptide was appended with an *N*-terminal glycine in position P-3 (G11-25) was also analyzed. If the variant including the OVA flanking

residues (QAV13-25) showed increased activity, it would decrease the likelihood that these flanking residues are specifically interacting with the DO11 TCR.

Interestingly, the QAV13-25 variant significantly increases HEL peptide activity, with the EC_{50} decreasing two-fold, from 6 to 3 μM peptide, while the maximum response increases two-fold, from 0.7 to 1.5 ng/mL IL-2. In contrast, the control peptides exhibit minimal or no change in activity (Figure 2.2; Table 2.2). The OVA flanking residues at positions P-4 through P-2 independently increases peptide activity in a manner independent of the specific T cell receptors involved. This supports a direct role for the QAV N-terminal flanking residues in stabilizing the peptide-MHC interaction.
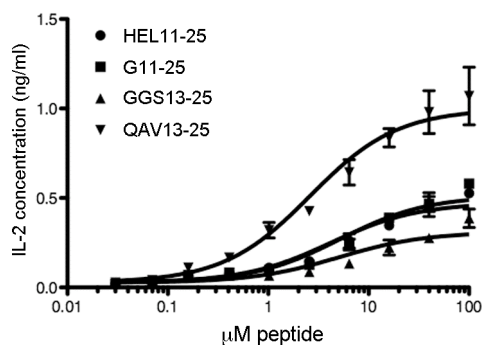
Figure 2.2. Ovalbumin flanking residues enhance HEL11-25 peptide activity.

HEL11-25 hybridoma cells and antigen presenting cells were incubated overnight in the presence of exogenous peptide variants, followed by quantification of IL-2 release as a measure of T cell activation. The wild type HEL11-25 peptide (●) is compared to a longer variant with an additional G at position P-4, G11-25 (■); an additional control variant in which the P-2 to P-4 flanking residues have been replaced with GGS, GGS13-25 (▲); and a variant presenting the ovalbumin- flanking residues in positions P-4 to P-2, QAV13-25 (▼).

2.3.4. Peptide activity is specific to wild type residues N-terminal to P-2

We thus sought to clarify the role of residues 325 and 326 in the register four interaction with I-A^d. We hypothesized that presentation of register four might actually be increased by a stable register three, mediated by rebinding or

sliding from the stronger register three to the weaker register four. While the more stable register three would be presented more frequently than register four, register shifting may provide a mechanism to capture the peptide before it diffuses away from the MHC after rapid register four dissociation. In this case, residues 325 and 326 would preferentially stabilize register three relative to register four, filling the canonical P-2 and P-1 positions for that register, versus an uncertain role in stabilizing register four.

To assess the possible role of alternate register stabilization, we first examined the requirement for specific residues in the flanking positions. Published crystal structures show conserved hydrogen bonds between main chain amino acids of presented peptides at positions P-1, with the side chain of histidine or tyrosine at $\beta 81$, and P-2, with the main chain of $\alpha 53$. In contrast, observed hydrogen bonds and van der Waals interactions between the MHC and more distant *N*-terminal flanking positions are rarer and more irregular, exhibiting residue specificity. We replaced the Q325 and A326 residues with glycine in a synthetic peptide (GGV328-337). If these residues simply form canonical main chain hydrogen bonds in the register three P-1 and P-2 positions then glycine substitutions should confer wild type activity. Glycine was selected over alanine because of the latter's propensity to occupy I-A^d anchor positions, as well as the presence of alanine at position 326 in the native epitope. The resulting variant GGV328-337 was more active than the 327-337 peptide (EC_{50} = 6 and 12 μM, respectively), but remained three-fold less potent than 325-337 (EC_{50} = 2 μM; Figure 2.1C, Table 2.2). Since wild-type peptide activity is modulated by Q325 or A326 or both, these residues may instead serve to stabilize register four relative to register three. For DO11, peptide activity is somewhat dependent on wild-type residues extending at the *N*-terminus to at least the register four P-3 position (residue 326).

To specifically probe the impact of register three stability on register four activity, we introduced two substitutions for valine at position 327 to modulate

43

register three stability. Among the candidate register three anchor residues (V327, A330, A332, N335), the P1 position occupied by V327 is the only anchor not overlapping with the register four core (329-337). Previously, residue substitutions in the spacious first binding pocket of I-A^d have been shown to have drastic impacts on register stability (107). We generated two peptide variants of GGV328-337: one in which the valine is replaced with a methionine (GGM328-337), which shows a strong preference in P1 (30), and a second in which the valine is replaced with serine (GGS328-337), which can act as an anchor, but poorly occupies P1 (100). If register four competes with register three for MHC binding, we would expect peptide activity to inversely correlate with P1 anchor preference, such that GGS > GGV > GGM. In fact, we observe a two-fold reduced EC$_{50}$ for GGS328-337, with similar activities for GGV and GGM (Figure 2.1D; Table 2.2); the role of register three remains ambiguous. Together with the substitutions in the HEL peptide, the flanking residue substitutions indicate that these residues stabilize register four, at least in part, in a manner independent of register three.

2.3.4. Chimeric Proteins Presenting Variant Peptides Recapitulate Synthetic Peptide Results

To confirm the peptide results and explore their relevance in the context of the native class II peptide display process, we introduced select peptide variants into a carrier protein, maltose binding protein (*malE*). If flanking residues are indeed involved in register four stabilization through MHC interaction, then their presence should be similarly required in fusion proteins and exogenous peptides. However, processing and subsequent presentation has been reported to at times eliminate display of weaker pMHC conformers observed during exogenous peptide display by means of the MHC class II-like endosomal protein H2-DM (105, 109-111). We wondered whether the presentation of register four might be

hindered by more stable competing registers when subjected to endosomal MHC loading conditions.

Peptide sequences were introduced into the exposed loop in place of residues 133-140 as first described by Martineau *et al.* (25), with a gly-ser extension (GSGSG) appended to the n-terminus of inserted epitopes to increase solubility (Figure 2.3A, Table 2.2). The activity of the *N*-terminally "truncated" peptides (325-337, 327-337, 328-337) maintains the relative preference for wild type flanking residues observed with synthetic peptides (Figure 2.3B, Table 2.2). The wild-type 323-337 and 325-337 inserts have EC_{50} within error (13 and 18 μM protein, respectively) while the activities of the 327-337 and 328-337 variants are greatly reduced (EC_{50} of 41 and 73 μM, respectively). The increased activity of variants retaining the P-3 and P-4 residues of register four during intracellular processing again supports their direct role in mediating the peptide and MHC interaction.
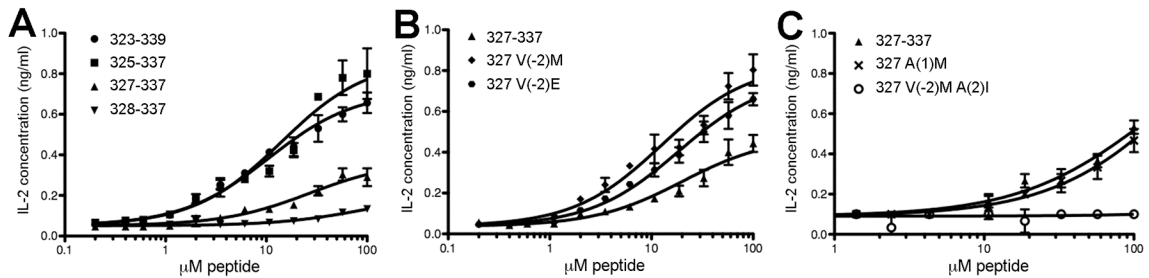
Figure 2.3. Specific flanking residues and register three enhance Ova peptide activity when endogenously processed as *malE* chimeras.

DO11 hybridoma and A20 antigen presenting cells were incubated overnight in the presence of *malE* chimeras containing ovalbumin epitope variants, followed by quantification of IL-2 release as a measure of T cell activation. *A,* Comparison of wild-type peptide 323-337 insert (●) with *N*-terminally truncated variants lacking alternate register anchor positions: 325-337 (■), 327-337 (▲), and 328-337 (▼). *B,* A truncated control peptide including the entire register four and register three, 327-337 (▲) was compared with variants in which the register three P1 anchor was altered in order to increase this

45

registers stability: 327 V(-2)M (◆) and 327 V(-2)E (⬟). *C,* The same control peptide 327-337 (▲) was compared with *malE* peptide variants in which the register four P1 anchor was altered from alanine to the more stable methionine, 327 A(1)M (✖) and a second variant in which the register three anchor positions P1 and P4 were both modified, 327 V(-2)M A(2)I (⭘).

2.3.5. Anchor substitution at position 327 enhances peptide activity within a carrier protein

Presentation of register four may be limited by the more stringent competition between a variety of epitopes for MHC binding during peptide loading. In general, greater concentrations of the carrier protein are needed to achieve equivalent stimulation as exogenous peptides (Table 2.2). In addition to the invariant chain peptide, which shows relatively high affinity for I-A^d (112), *malE* itself contains several immunogenic determinants with which the inserted epitope must compete (113). We hypothesized that under these conditions, we might observe register four dependence on binding of register three. The experiments were designed carefully to ensure that any result from the distinct processed epitopes is not explicable by differences in solvent exposure or protease sensitivity.

To test the relationship between register three stability and epitope activity, we therefore produced two constructs based on *malE* 327 with strong P1 anchors: 327 V(-2)M and 327 V(-2)E. The negatively charged glutamic acid side chain contrasts with the nonpolar methionine in potential contributions to solubility and protease accessibility, but both are strongly preferred in the P1 binding pocket (107). Both substitutions enhanced activity of the processed epitope about two-fold, based on EC$_{50}$ values (18 and 25 μM chimera versus 41 μM for the native 327-337 fusion protein; Figure 2.3C), consistent with a weak effect of register three stabilization on activity.

To probe the extent of this observed effect, we create an additional mutant to stabilize register three, 327 V(-2)M, A(2)I. In this construct, residue four of the

46

third register was modified from alanine to the anchor residue-preferred isoleucine in addition to the previously reported methionine substitution at anchor position one. We also attempted to stabilize the active register four in the same fashion as register three, by substituting methionine at anchor position one to create 327 A(1)M. The single register four anchor variant, 327 A(1)M, did not significantly impact activity, while the double register three variant 327 V(-2)M, A(2)I eliminated activity (Figure 2.3D). This is consistent with our earlier observations with synthetic peptides that anchor residue changes within the register four core may compromise T cell recognition.

2.3.6. Differences in chimera versus peptide activity are not dependent on H2-DM

During peptide loading onto the MHC in the endosomal compartment, the class II-like H2-DM protein modulates the repertoire of displayed peptides (32, 34). To examine whether the differential activity of our chimeras is a result of H2-DM activity, we substituted 3A5, a variant of the A20 cell line with an inactivated H2-DMα chain (103, 104) in our stimulation assays. While these experiments were less sensitive, the relative activities of the peptide variants were consistent with A20 experiments for the peptide length controls (Figure 2.4A), and the 327 anchor substitutions (Figure 2.4B).
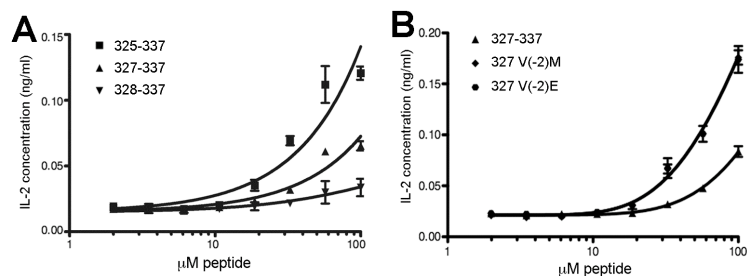
Figure 2.4. H2-DMα is not responsible for the different activities of endogenously processed peptide variants.

DO11.10 hybridoma and 3A5 antigen presenting cells containing mutant H2-DMα gene, were incubated overnight in the presence of *malE* chimeras containing ovalbumin epitope variants, followed by quantification of IL-2 release as a measure of T cell activation. *A*, To assess the effects of removing alternate register anchor positions, truncated ovalbumin peptide inserts 325-337 (■), 327-337 (▲), and 328-337 (▼) were compared. *B*, To assess the effect of stabilizing register three with P1 anchor substitutions, the control peptide 327-337 (▲) was compared with variants 327 V(-2)M (♦) and 327 V(-2)E (●), substitutions which have been shown to confer enhanced I-A^d binding. Curve fits for the latter two are nearly identical, and lie atop one another.

The use of the *malE* carrier protein to mediate peptide presentation supports the role of flanking residues in stabilizing register four presentation, with sequence preferences maintained upon intracellular processing in a H2-DM-independent manner. This supports a role for flanking residues in interacting with the MHC, either through stabilizing the recognized conformer or an intermediate binding state. Additionally, we find that the processed epitope is more effectively presented when register three is stabilized, an effect not seen using exogenous peptides.

2.3.7. Predicted peptide-MHC interactions

To further probe the relationship of the binding registers and flanking residue interactions with the MHC, we used the molecular docking program ClusPro in conjunction with the experimental I-A^d structure to predict peptide-MHC interactions (106). Subsequent docking of the OVA peptide yielded high scoring structures with the peptide oriented in register one as seen in the crystal structure, and register three, but neither registers two or four (Figure 2.5).
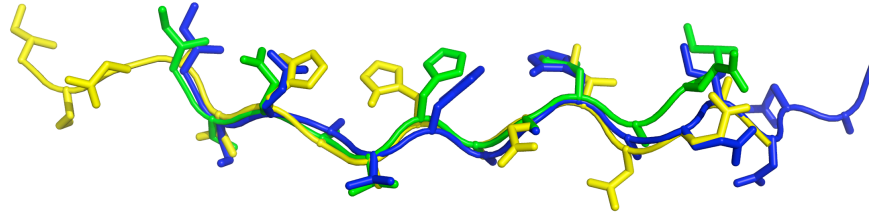
Figure 2.5. Ovalbumin registers one and three are predicted as stably associated with I-A^d MHC by computational docking.

Peptide and MHC coordinates of 323-339 and I-A^d were extracted from PDB 1IAO and computationally docked using Cluspro (106). Shown are peptide alignments of the register one wild type crystal structure (green) with the high-scoring ClusPro predicted register one alignment (blue), and ClusPro predicted register three alignment (yellow). Registers two and four were not predicted.

2.4. Discussion

Register masking has attracted increasing attention in recent years as a mechanism for promoting autoimmunity via incomplete thymic education. For instance, the weakly bound myelin basic protein epitope (MBP ^{Ac}1-9) is flanked at the *c*-terminus by a more stable register (91, 92), as well as at the N-terminus in the golli-form of MBP expressed in the thymus (93). Maverakis *et al*. observed direct competition between registers for MHC binding, as presentation of each register was reduced by the presence of an alternate register. In the non-obese diabetic mouse model of type 1 diabetes, the insulin B 9-23 peptide was also found to bind in multiple registers (114, 115), with the least stable register apparently activating pathogenic T cells (90). By measuring DO11 T cell responses in the presence of a series of ovalbumin peptide variants, we

demonstrate that the weakly bound fourth register of the ovalbumin 323-339 peptide is not masked by the two overlapping, more stable registers. In fact, responses are maximized when both registers three and four are present. This contrasts with the current view of register masking and we thus aimed to understand the roles of alternate registers in DO11 activation as a potentially novel mechanism to enhance and stabilize the display of weakly bound peptides.

Of the four registers present within the ovalubumin 323-339 peptide, the first register (residues 324-332), forms the most kinetically stable complex with MHC (95), is preferentially presented in I-A^d tetramers (97), and in the I-A^d crystal structure (100). Ovalbumin peptides with *N*-terminal truncations to position 325 no longer occupy the first register, due to the absence of the register one P1 anchor and P-1 flanking residue (95). Yet, in the presence of the responsive DO11 hybridoma, the resulting 325-337 peptide is no more active than the full-length version, indicating that register one does not competitively inhibit display of the active register four. In contrast, the presence of an intact register three enhances peptide activity relative to the isolated register four: the activity of the short 328-337 peptide is lower than that of the 327-337 peptide, which is in turn lower than the 325-337 peptide. Our efforts to alter register three stability through anchor one substitutions showed no clear effect for exogenous peptide, while activity did correlate with predicted stability for epitopes inserted into the *malE* carrier protein. The latter result occurs in a system designed to replicate natural protein processing, and seems to indicate a "piggybacking effect," in which display of a weaker binding register is enhanced by the presence of a stronger, overlapping register. Since activation requires recognition of peptide by both the MHC and the TCR, a register weakly bound by MHC might be immunodominant if it presents a composite peptide-MHC surface readily recognized by a variety of T cell receptors. Indeed, our findings are consistent with the surprising finding of Landais *et al.* that the weakly bound register four is the immunodominant register (97).

However, register four activity does not appear to be entirely dependent on the piggybacking of register three. Our data suggests that previous efforts to measure register four stability, in which the peptide was truncated at the register four P-2 position, with P-2 itself altered from a valine to a lysine to prevent register shifting (327-339 KIV) (99), underestimated the value for the native register. Moreover, substitution of the register four P-4 and P-2 residues into the distinct I-A^d epitope HEL11-25 dramatically increases peptide activity when presented to a responsive T cell hybridoma. Within the available peptide/MHC crystal structures, flanking residues typically form hydrogen bonds and van der Waals contacts with the MHC, which would stabilize weak registers. These interactions are commonly seen out to P-2, and occasionally to P-3 and beyond (116-122). In the ovalbumin case, the register four P-4 to P-2 flanking residues are also employed as the register three P-2 to P1 residues (Table 2.1). Thus, efforts to measure the stabilizing effects of these residues on register four will be obscured by the presence of the intact register three with a slower dissociation rate. In summary, analysis of synthetic peptides shows no clear correlation between register four activity and register three stability. This may be due the reduced hybridoma sensitivity to peptides versus endosomally processed protein, increased register three stabilization eventually negatively competing with register four or the flanking residues stabilizing register four via a register three-independent mechanism.

An alternative explanation for the apparently co-operative interactions of registers three and four is that they form a single, larger register as opposed to two distinct, yet overlapping registers. Several lines of evidence refute this hypothesis. First, McFarland and Beeson convincingly demonstrated the existence of register three through the use of an amino acid substitution at P2 (H328Q), which forms a dissociation intermediate with the MHC. Second, use of the ClusPro docking program to predict interactions between the ovalbumin 323-339 peptide and empty I-A^d molecules rank register three as highly stable (Figure

2.5). Third, register four has been independently proposed multiple times: by the Marrack-Kappler group (personal communication), Robertson *et al.,* and Landais *et al.* In the latter report, I-A^d tetramers covalently linked to the register four nine amino acid core, but not those linked to register three, bound and activated DO11 T cell hybridomas. In our hands, the free peptide consisting of the nine amino acid core was sufficient to activate DO11, albeit at extremely high concentrations (data not shown).

Our finding that the anchor substituted peptide 323-339 EIVS severely diminished activity contrasts with the finding of Robertson *et al*. Many reports have described successful increases in biological activity via anchor swaps, particularly for CD8$^+$ T Cells (123-129). However, the effect seems far from general, as several other reports have shown the opposite effect include (116, 130-132) Notably, our work is consistent with Landais *et al*., who generated I-A^d tetramers with the register four 329-337 peptide and observed reduced DO11 staining when the chemically distinct P1 substitutions A329M or A329E were introduced (97). The discrepancy with Robertson *et al.* may be attributable to assay differences: measurement of hybridoma IL-2 secretion versus their DO11 transgenic T cell proliferation. In summary, this result highlights the sensitivity of the DO11 TCR to even small conformational changes introduced within the core.

A key outstanding question is the mechanism by which stronger registers competitively inhibit display of the masked register in myelin basic protein (93), yet appear to co-operatively enhance display of ovalbumin register four. One possibility is that the relative stabilities of the ovalbumin registers are similar, and the registers highly overlapped, such that the energetic costs of peptide transfer between registers three and four are small. This sliding or rebinding mechanism could prevent peptide loss due to rapid register four dissociation and subsequent peptide diffusion away from the antigen-presenting cell. Alternatively, the I-A^d molecule itself, with its small, hydrophobic anchor pockets, may be more permissive to register shifting.

Overall, our findings support inclusion of residues outside the nonameric core and multiple, distinct registers in epitope prediction models (79, 80, 85, 86). The ability of ovalbumin flanking residues to confer enhanced activity to an unrelated HEL11-25 peptide may have general applications for understanding peptide immunogenicity and stability. In terms of the specific ovalbumin 323-339 peptide, our attempts to increase peptide activity through register four anchor substitutions severely diminished activity. Anchor substituted peptides may be more prone to influencing T cell response, as compared to flanking residue changes which are more distal to primary T cell contacts. In contrast, modification of flanking residues may increase activity of a weakly bound peptide while minimizing changes in the responding T cell repertoire. Recently described high throughput techniques (133), are likely to be instrumental in identifying stabilized variants of weakly bound peptides for use as altered peptide ligands in autoimmune therapies and vaccines.

CHAPTER 3

T CELL RECEPTOR ENGINEERING

3.1. Introduction

The primary goal of this work was the creation of a soluble, high-affinity version of the DO11 T cell receptor for the detection of ovalbumin displayed by the murine MHC, I-A^d. Secondary focus was given to the manipulation of expression and display systems toward generality in the display of other T cell receptors. Consideration was also given to tethering of immunoactive peptides to soluble constructs.

3.1.1. Single-Chain T Cell Receptors

The creation of T cell receptor reagents requires modification from their native membrane-bound form. This modification has taken several forms including the single chain T cell Receptor (scTCR), the disulphide-linked T cell Receptor (dsTCR), and c-jun/v-fos leucine zipper-linked TCRs (Figure 3.1). A single-chain TCR is made up of the smallest functional unit of the T cell receptor, its α and β chain variable regions connected by a flexible linker. The scTCR mimics a popular antibody creation, the scFv, which has been applied extensively in the development of antibody therapeutics. The extracellular regions of membrane-bound TCRs are structurally homologous to antibody fragments, comprised of two chains and four immunoglobulin domains (two constant, two variable, with the variable (V) regions responsible for engaging ligand. The first scTCR was created by Novotny *et al.* (134), and many have since been attempted.
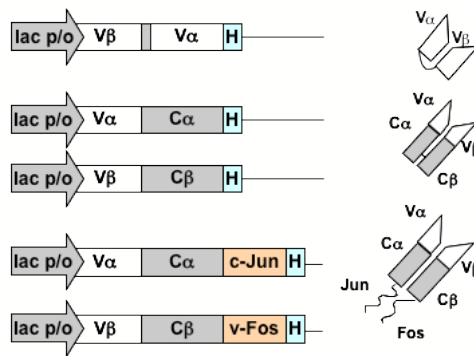
Figure 3.1. T Cell Receptor Expression Constructs.

Previously, DO11 has been expressed solubly in insect cells by a construct consisting of full DO11 chains (α and β chains with variable and constant regions) truncated just prior to the transmembrane portion of the constant domain (135). The dsTCR follows this general pattern, but introduces a non-native disulphide bond introduced between constant regions (136). Others have opted for the inclusion of a leucine-zipper region rather than a disulphide bond (137). The interaction between these chains has been reported to stabilize the molecule. While both the leucine zipper and dsTCR constructs produce large quantities of soluble TCRs, both also rely on the refolding of aggregated intracellular inclusion bodies when expressed in bacteria. Inclusion bodies are insoluble accumulations of improperly folded protein in the cell, and correlate negatively with the phage display of fusion proteins.

Described here is the production of scTCRs solubly in bacterial cytoplasm followed by export to the periplasm and release upon osmotic shock. Soluble expression in the robust, fast-growing bacteria is inexpensive and allows for relatively rapid evaluation of mutant constructs. This work shows that this method is far from generalizable, but highly useful in specific instances, as is the case for DO11.

3.1.2. Multimerization and Affinity Maturation

The utility of scTCRs as reagents is dependent upon overcoming inherent low-affinity binding to substrates, typically between 5-100 μM (138), which is relatively weak in comparison to the low to sub-nanomolar affinities routinely found in antibodies. Multimerization employs multivalent constructs to increase the overall avidity of a reagent, whereas affinity maturation is seeks to directly increase the affinity of the monomer.

Affinity is quantified as the dissociation constant of the TCR/pMHC complex, $K_d = k_{off}/k_{on}$.

$$\text{TCR} + \text{pMHC} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{TCR-pMHC}$$
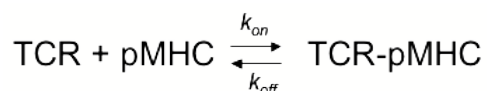
Affinity maturation is routinely achieved for scFvs, but work with soluble TCRs has been limited for the most part to work done by two groups. Kranz and colleagues have worked extensively with a yeast display system in engineering the 2C class I system, and more recently, the 3.L2 class II system (139). Yeast display was used to select for thermally stable variants of the 2C scTCR (140) and improve its affinity 100 fold, to 9 nM (141). These higher affinity mutants, however, showed high cross reactivity to self-peptide/ MHC (142). Later work did indeed generate high affinity, highly- specific TCRs been generated through yeast display, for the 2C (143) and 3.L2 scTCRs (139). Recently, an scTCR based on the 2C scaffold was shown to be affinity-matured through display on a hybridoma via retroviral infection (144). The maturation of 3.L2 represents the only instance of maturation towards a Class II pMHC target of which we are aware.

Phage display of only the Vα domain of a TCR was originally reported in 1995 (145), with display of a functional TCR was first reported shortly thereafter by Weidanz et al. (146). These authors displayed DO11 as three-domain scTCR (including the constant β region), and demonstrated activity, but no affinity maturation was reported. Since then, affinity maturation via phage display has

been reported for three TCRs. For A6, expressed as a dsTCR (147), affinity was improved from about 32 μM to 26 pM against a viral Class I pMHC, and the matured TCR showed no cross-reactivity to non-specific peptides. Of note, the A6 TCR was also later affinity matured through rational design based on the known crystal structure (148). Maturation of the 1G4 dsTCR against a tumor-associated class I antigen was later shown strictly through the introduction of CDR2 mutants (149). Despite the canonical association of CDR2 with the MHC, variants were shown to be peptide-specific. A third TCR, specific to the class I-restricted SL9 antigen of HIV was isolated from a T cell line and affinity matured against its target through CDR2β and CDR3α mutations, giving an increase from a wild type K_d of 143 nM to less than 400 pM.

Multimerization is typically accomplished through the natural interaction of four biotins with a single avidin variant. MHC tetramers, in which the c-termini of each MHC molecule is singly biotinylated followed by complexation with avidin, have been utilized so extensively in characterization of T cell specificity that an entire NIH facility exists to manufacture these reagents (for review: (150)). Tetramers have been created previously for TCRs in the context of dsTCRs (151) and leucine-zipper heterodimers (152). Both reports showed an effective increase in TCR avidity as a result of multimerization. Laugel *et al.* demonstrated that tetramers of the A6 TCR can compete with cytotoxic lymphocytes for binding to antigen presenting cells *in vitro*. Both of these cases utilized TCRs with relatively high affinity to mitigate the affects of naturally low levels of antigen display.

In this work we demonstrate the affinity maturation of DO11 against KJ1-26 through the use of phage display. The clonotypic antibody KJ1-26 was used as a proxy for ova 323-339/I-Ad constructs because of previously documented issues with the soluble, covalently-linked pMHC reagent. This antibody was first isolated as binding DO11 (153), and later shown to isolate a single distinct T cell hybridoma from a group of 397 hybridomas isolated from BALB/c T cells (154).

The isolated hybridoma, 7DO-286.2 was shown to share fine specificity with DO11. Given this specificity for the binding site of DO11, we hypothesized that it mimics the pMHC surface and could serve as a pMHC proxy during DO11 affinity maturation by directed evolution.

After maturation, we demonstrate multimerization of the scTCR and its matured variant through natural binding of a monoclonal antibody to an EE epitope tag inserted into the construct. The resulting dimer was then further complexed to a fluorescently labeled IgG1, and used to stain A20 cells. The DO variant "176" showed labeling activity based on the presence of 323-339 peptide.

3.1.3. Therapeutic Applications

High affinity variants of DO11 have potential for in further exploring the antigen presentation of ovalbumin by I-Ad. In addition, affinity matured TCRs as a class may be useful in diagnostics and treatment of cancer, viral infection, and autoimmunity either in soluble form (Reviewed: (155), (156)), or transduced into the T cells of an individual patient (157). Class II reactive TCRs are uniquely suited for targeting antigen-presenting cells in autoimmune conditions, which are mediated by CD4$^+$ T cells. The multitude of issues involved in such a treatment (necessary ligand density, peptide specificity of the reagent, immunogenicity of the construct) are beyond the scope of this work. We do, however, demonstrate that the fusion of a human IL10 (IT9302) mimic peptide shown to possess many of the immunosuppressive functions of the cytokine (158-160) can be produced as a conjugate to a 172 TCR in soluble bacterial expression at high levels (>30 mg/L). Cytokine tethering would seem ideally suited for TCR targeting, based on the association of both with antigen-presenting cell based receptors, and potential side effects associated with global circulation.

3.2. Materials and Methods

3.2.1. Single Chain T Cell Receptors

Full chain T Cell receptor genetic material was obtained from various sources. 2B4, 2C, 172.10, 1934.4 G206, and D10 were gifts from KC Garcia, Stanford University. 2D1 was a gift from Lars Fugger, Cambridge University. The 3B6 and 4C4 T Cell Lines were gifts from Thomas Forsthuber, University of Texas- San Antonio. Lymph Nodes from a DO11 transgenic mouse were a gift of Marc Jenkins, University of Minnesota. When necessary, total RNA was isolated from relevant cells using Invitrogen's *PureLink Micro to Midi Total RNA Purification System* (12183-018, Carlsbad, CA), or Qiagen's *RNeasy Mini Kit* (#74104, Gaithersburg, MD). Alpha and beta variable regions were connected via a $(Gly_4Ser)_4$ linker using overlap PCR and gene-specific oligonucleotides. Fusions of various peptides to the N- and c- terminus of scTCR constructs were synthesized using short oligonucleotides (<60 bp) and incorporated into scTCRs using overlap PCR.

After amplification using flanking primers incorporating directional SfII restriction sites, scTCR constructs were PCR purified (Qiagen #28104), digested with SfiI (New England Biolabs #R0123, Ipswich, MA) at 50 °C for greater than 3 hours, separated on a 0.8% agarose gel and purified using a QIAquick Gel Extraction Kit (#28704). Purified scTCR DNA was then mixed with similarly purified expression plasmid at between three and five to one molar ratio and ligated using New England Biolabs' T4 DNA Ligase (#M0202) at 16 °C for greater than one hour. Following transformation into XL1B *E. coli* strain, plasmids were isolated from individual colonies using a QIAprep Spin Miniprep Kit (#27106) and sequenced to confirm incorporation.

3.2.2. Plasmids and Phagemids

Plasmids pAK400 and pAK500 were a gift from Andreas Pluckthun, University of Zurich. The phagemid pMoPac24 was a gift from George Georgiou, The University of Texas at Austin.

Phagemid pMoPac24 was modified to express variants of M13 Gene Product VIII (pVIII) in the place of gene product III (pIII). Gene VIII fusion protein and it's variants were synthesized by amplification of 50-60 bp overlapping oligonucleotides with flanking primers incorporating 5' Asc I and 3' HindIII restriction sites. As with the original gene PCR products were purified, digested at 37 °C for than 1 hour, separated on a 2% agarose gel and gel purified. Following transformation in XL1B *E. coli*, plasmids were isolated from individual colonies and sequenced to confirm incorporation.

Plasmid pAK400 was modified to express peptide fusion peptide IT9302 through by full plasmid amplification. The plasmid was methylated using CpG methyltransferase (NEB #M0226) at 37 °C for greater than one hour. Oligonucleotides incorporating the insertion with at least 18 bp showing complementarity to the site of insertion were phosphorylated using T4 polynucleotide kinase (NEB #M0201) for at least one hour at 37 °C. Phosphorylated primers were then used to amplify the plasmid in PCR. Following amplification, the original plasmid was digested with DpnI (NEB #R0176) at 37 C, and the amplified plasmid ligated using T4 DNA Ligase (NEB #M0202) prior to transformation in the XL1B strain of *E. coli*.

3.2.3. Protein Expression

Sequenced plasmids containing scTCR genes were transformed into the strain BL21, grown overnight at 25°C in baffled shake flasks ranging from 500 mL to 4 L. After replacing spent media with fresh terrific broth, protein expression was induced with 1 mM IPTG for 4-24 hours at 25 °C. The cells were harvested by centrifugation and cell pellets were resuspended using ice cold 0.1 M Tris pH

7.4-buffered 0.75 M sucrose at 3 mL per gram of cell pellet. An osmotic shock to release the periplasmic contents was performed by addition or 3 mL 1.0 mM EDTA per gram of cell pellet and approximately 0.3 mg/mL lysozyme followed by the addition of 6 mM $MgCl_2$ on ice. After an additional centrifugation step, the supernatant was dialyzed against 10 mM Tris pH 7.4, 0.5 M NaCl with 12-14 kDa MWCO dialysis tubing. Chimeric proteins were recovered via immobilized metal affinity chromatography (IMAC) with the c-terminal His tag and size exclusion chromatography (Superdex 75 or 200; GE Healthcare) in HBS and stored at -80 °C prior to use.

3.2.4. Biotinylation of scTCRs

Purified scTCR tagged with a biotinylation site was concentrated to at least 40 μM using centrifugal concentration. Proteins were biotinylated using biotin protein ligase (BirA, Avidity, Denver, CO) by incubating overnight at room temperature. Protein was then purified by size exclusion chromatography (Superdex 75 or 200; GE Healthcare) in HBS and concentrated. Biotinylation was evaluated by mixing with equimolar extravidin (Sigma #E2511), separating on 4-20% SDS-Page gradient gel, and detecting via α-His HRP western blot.

3.2.5. Preparation of EE-Tagged Dimers and Tetramers

Epitope tags were inserted into specific proteins using overlap PCR as described above. When necessary, proteins were concentrated to 3-5 mg/mL using a centrifugal system. Proteins were quantified using a commercial bicinchoninic acid assay (Pierce #23227, Rockford, IL) and by absorbance at 280 nm in conjunction with a theoretical extinction coefficient. Dimers were prepared from proteins containing the "EE" epitope tag (amino acid sequence EYMPME) by stepwise addition of the monoclonal antibody specific to that tag (Covance #MMS-115P, Princeton, New Jersey). Monoclonal antibody was added to a 1:2 molar ratio in a series of 6-10 individual steps, separated by 10 minutes. Tetramers were prepared by addition of α−IgG1 monoclonal antibody (clone x56,

BD Pharmingen #550874) specific for α-EE, fluorescently labeled with allophycocyanin (APC). This antibody was added to dimerized proteins in an additional series of stepwise additions at a 1:2 molar ratio to the initial antibody.

3.2.6. Flow Cytometry

A20 cells expressing I-A^d were incubated with specific doses tetramerized scTCR or maltose binding protein control along with specific doses of peptide. Incubation was carried out at 4 °C in in Hepes buffered saline with 1% FBS for 2-4 hours. Allophycocyanin staining of cells was measured using an LSR II Fortessa flow cytometer (BD Biosciences), stimulated at 635 nm, with fluorescence measured after emitted light passed through a 670/30 band-pass filter. Data were analyzed using FlowJo 9.4.9 (Tree Star, Inc.; Ashland, OR).

3.2.7. Phage Production

XL1-B cells containing phagemid were grown in baffled erlenmeyer flasks at 37 °C in the presence of 200 μg/mL ampicillin and 10 μg/mL tetracycline to 0.5 OD600, 25 mL to 1L total volume. M13KO7 was added at a multiplicity of infection (MOI) of greater than one along with 1 mM IPTG and cultures continued to shake for 2 hours at 37 °C. Kanamycin was then added (50 mg/mL), and culture temperatures were lowered to 25 °C, where they were allowed to shake for up to 24 hours before harvest.

3.2.8. Phage Purification

Supernatants from XL1-B cultures were collected from cultures were collected following 3,000xg centrifugation for 30 minutes. Phage were precipitated by the addition of one fifth volume of 2.5 M NaCl, 20% w/v PEG 8000 to supernatants, followed by 1 hour incubation at 4 °C. Phage were collected by after centrifugation of supernatants (3000xg for at least one hour), and resuspended in 1 to 2 mL HBS. Resuspended phage were then centrifuged in at 16,000xg for a minimum of 10 minutes to remove subcellular debris. Phage

were then precipitated again with 1/5 volume 2.5 M NaCl, 20 % w/v PEG 8000 and resuspended in 0.5 – 1 mL HBS.

3.2.9. Library Construction

A library of DO11 CDR3α variants was produced according to the protocol of Weiss *et al.* (72). DO11 scTCR in pMoPac24 was modified to include three consecutive stop codons starting at amino acid position 107 (IMGT numbering) by quick change PCR, as previously described.

Two libraries were constructed to produce libraries of 8 and 10 amino acid CDR3α lengths. Oligonucleotides were synthesized to randomize six codons of DO11 CDR3α, from either amino acid 107 through 115 (ten amino acid library) or 106 through 116 (eight amino acid library). Each oligo included 18 nucleotides flanking the random codons that were identical to wild type DO11. Random codons were coded as NNS, where "N" corresponds to any of the four-nucleotide bases and "S" is guanine or cytosine.

Single-stranded DNA template was isolated from phage particles. CJ236 *E. coli* transformed with pMoPac24 containing DO11 with stop codons in its CDR3α was grown in 1 mL of terrific broth containing 0.25 μg/mL uridine, 34 μg/mL chloramphenicol to select for the F' episome, 200 μg/mL ampicillin to select for the phagemid, and 1% glucose. When the cells reached 0.5 OD_{600}, they were infected with M13K07 helper phage, continued to grow for 1 hour, and scaled up to 30 mL terrific broth containing 50 μg/mL kanamycin. Cells were allowed grown overnight at 37 °C with 200 rpm shaking. Phage were harvested as before, precipitated, and resuspended in 0.5 mL PBS. Approximately 20 μg of single-stranded DNA was purified from this phage using a QIAprep M13 Spin Kit (#27704). DNA produced from CJ236 contains a significant amount of uracil in place of thymine.

The library was then created from by extension around a dU-ssDNA template with randomized oligonucleotides. Oligos were phosphorylated with T4

polynucleotide kinase, and mixed with template in approximately a 3:1 molar ratio. The mixture was heated to 90 °C in a thermocycler for 2 minutes, cooled to 20 °C over the course of four minutes, and allowed to incubate at five minutes at 20 °C. The annealed oligo was extended around the plasmid using T7 DNA polymerase and ligated with T4 DNA ligase while incubating overnight. The resulting template was then transformed into electrocompetent XL1-B, which preferentially replicates the newly synthesized strand lacking uridine.

3.2.10. Affinity Maturation

XL1-B cells containing DO11 CDR3α library were grown up and phage were produced and purified as described above (3.2.6 and 3.2.7). Between 10^8 and 10^{10} cfu were added to immobilized KJ1-26 or irrelevant antibody in 96-well plates and allowed to incubate for 2 hours at 37 °C with gentle shaking. Cells were then washed five times with PBS and eluted with 50 mM glycine pH 2.2. Eluted phage were rapidly neutralized with 2 M Tris, and infected into 2 mL ER2738 cells at 0.5 OD$_{600}$ growing in 2xYT with 10 μg/mL tetracycline for 30 minutes at 37 °C with shaking. Following infection, these cells were scaled up to 100-500mL TB containing 200 μg/mL ampicillin and 50 μg/mL kanamycin, 1 mM IPTG, and 0.1% glucose. This media lacked tetracycline, to prevent cross infection of library variants. Cells were shaken at 37 °C until approximately 0.5 OD500, then cooled to 25 °C and permitted to shake overnight. The following day, phage was purified and another round of panning performed.

3.2.11. Antibodies and Peptides

Synthetic ovalbumin 323-339 peptide (p323) was purchased from Anaspec (#62571, Fremont, CA). α−M13 HRP was purchased from GE Healthcare (#27-9421-01). KJ1-26 was purchased from BD Pharmingen (#551771). α-his HRP was purchased from Invitrogen (R931-25, Carlsbad, CA).

3.2.12. Oligonucleotides

All oligonucleotides were purchased from IDT (Coralville, Iowa).

3.2.13. Antigen presenting and T cell hybridoma cell lines

The murine A20 B cell lymphoma line presenting I-A^d MHC was acquired from ATCC (TIB-208). The murine DO11 T cell hybridoma line was transferred from Dr. Philippa Marrack, National Jewish Health Center, Denver, Co. All cells were maintained in DMEM supplemented with 10% FBS, 50 Units penicillin/ 50 mg streptomycin per mL and 2 mM L-glutamine.

3.2.14. Inhibition of T Cell Activation

Each well of a 96-well plate was seeded with 15,000 A20 cells, 15,000 T cell hydridoma cells, 20 μM ovalbumin peptide, and specified concentration of M13 bacteriophage wild type or displaying specified proteins in each well with DMEM media without L-glutamine. After a 24-hour incubation at 37 °C, 5% CO_2, supernatant was transferred to a new plate and stored at -80 °C prior to IL-2 quantification assays. All phage concentration doses were performed in triplicate wells and all experiments replicated at least twice. In each well, the volume fraction of phage in HBS solution added did not exceed 20%, and an FBS concentration of 10% was maintained in all wells. The IL-2 concentration in the supernatant was quantified using a matched pair ELISA quantified with an IL-2 standard as described (BD #555148). Data was analyzed with Graphpad Prism 5 (La Jolla, CA).

Glutamine was excluded from experiments (except where noted) to minimize cell growth, as preferential growth was observed in inverse proportion to T cell stimulation. Since overall IL-2 release is effected by the ratio of T cells to B cells, and assays were performed only at a limited number of discrete doses, Ireasoned that differential growth could potentially mask small differences in the activity of specific peptides. The absence of glutamine lowered overall IL-2

release and required larger concentrations of peptide or protein to reach a given release.

3.2.15. Assessment of scTCR Binding Affinity Using BIACore

Kinetic binding assays were performed with scTCRs and KJ1-26 using a BIAcore 2000 (GE Healthcare). KJ1-26 were coupled to CM5 chips using NHS-EDC chemistry to a level of ~600 RU. The signal from a flow cell coupled with a control protein (BSA) was used to correct for non-specific binding to the matrix. scTCRs were injected in a duplicate two-fold dilution series from 400 nM at a flow rate of 35 μl/min. The association rate constant (k_{on}), dissociation rate constant (k_{off}), and equilibrium dissociation constant (K_d; $K_d = k_{off}/k_{on}$) were calculated assuming a Langmuir 1:1 binding model with BIAevaluation software.

3.3. Results

3.3.1. Cloning of Single Chain T Cell Receptors

Single chain T cell receptors were assembled from genetic material of various origins. Variable regions were isolated and joined on either side of a flexible linker made up of repeating gly-ser subunits. The resulting scTCRs were cloned into application-specific members of the pAK vector series. These vectors were originally developed in the Pluckthun lab (65), and offer convenient shuttling of genes between gene III phage display on M13 and periplasmic protein expression driven by a lacZ promoter. The two primary vectors used in this work are pAK400 and pMoPac24. For the purposes of periplasmic expression of soluble, pAK400 connects the inserted gene to a pelB leader sequence at the *N*-terminus and a hexa-histidine tag (6xHis) at the c-terminus via a directional Sfi I sites (Figure 3.2). Translated proteins are transported into the periplasm via the pelB leader, where they are cleaved and released following osmotic shock. The histidine tag then allows for convenient purification via affinity to divalent cations, also known as immobilized metal affinity chromatography. pMoPac24 was

derived from the pAK400 vector by Hayhurst *et al.* (161), who modified it by connecting the inserted gene to a short version of M13 gene III (isolated from pAK100), coexpressing the seventeen kilodalton protein (Skp), and substituting ampicillin for chloramphenicol resistance. Ampicillin resistance is useful in conjunction with the chloramphenicol-resistant CJ236 cell line used in library generation.
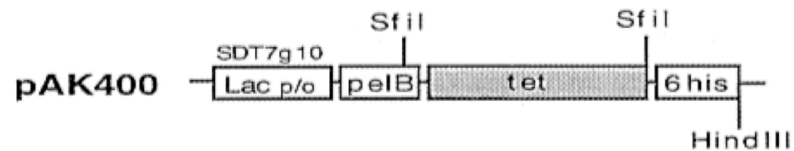
Figure 3.2. pAK400 Schematic.

ScTCRs to be expressed were cloned in place of the tet cassette with directional SfiI restriction sites. (Source: (102)).

A representative list of the scTCRs studied for this work is shown in Table 3.1. Receptors such as DO11, 2B4, and D10 were among the first hybridoma models created from mouse T cells, and continue to be well studied as mouse models. Others are of particular interest in autoimmunity, having been isolated as autoreactive T cells in particular mouse models of type I diabetes (G206), and multiple sclerosis (172.10, 1934.4, 4C4, 3B6). DO11 is of primary interest due to its ovalbumin target, while the remainder are included to illustrate the generality (or lack thereof) of the documented platforms.

TCR	Origin	Type	TRAV	TRBV	Peptide	MHC	Expression (mg/L)	References
DO11	M	CD4	5D-4	13-2	OVA 323-339	I-Ad	30-50	White 1983 J. Imm.
172.10	M	CD4	14-3	13-2	mMBP Ac 1-11	I-Au	30-50	Wraith 1989 Cell
1934.4	M	CD4	16D-7	13-2	mMBP 1-23	I-Au	10	Urban 1988 Cell
D10	M	CD4	14D-2	13-2	ConA 134-146	I-Ak	30	Kaye 1983 JEM
G206	M	CD4	3-3	13-2	GAD65 206-220	I-Ag7	10	Kim 2004 PNAS
4C4	M*	CD4	7D-2	4	hMBP 82-100	DRB5-*0101	1	Kawamura 2008 Jim
3B6	M*	CD4	7D-2	20	hMBP 82-100	DRB5-*0101	1	Kawamura 2008 Jim
2B4	M	CD4	4-2	26	PCC 81-104	I-Ek	1	Hedrick 1982 Cell
2D1	H	CD8	12-2	24	Human PLP 45-53	HLA-A3	3	Honma 1997 JNeurIm

*T cell lines raised in mouse against human transgenic MHC

Table 3.1. T Cell Receptor Origins, Antigens, and Single Chain Expression Level.

3.3.2. Protein Expression

3.3.2.1. TRBV13-2 Associated Stability

Single-chain T cell receptors in pAK400 were transformed into the *E. coli* strain BL21, which is commonly used for protein expression due to its rapid growth characteristics and deficiency for the proteases *ompT and lon*. Cells were typically induced for between 12 and 20 hours prior to harvest. Proteins were isolated by osmotic shock and purified by binding of the histidine tag to Ni^{2+} associated sepharose beads and subsequent FPLC size exclusion. The expression of various scTCR constructs is shown in Table 3.1. In general, we found high-level protein expression is associated with receptors incorporating the TRBV13 (alternately known as Vβ8) variable region. This expression is dependent upon both the orientation of the chain with respect to its associated alpha chain.

Four amino acid changes have previously been reported to stabilize the TRBV13-2 chain. The G17E mutation was selected in for in a randomly generated yeast display designed to improve expression of the 2C scTCR (162). Similarly, the H54Y mutant was selected in 2C from a yeast display library subjected to thermal stress (140). Two other mutants, I90T and L93S, were reported as having been originally identified to improve expression in *E. coli* (163). All amino acid numbering is according to IMGT standards. The incorporation these four substitutions into the D10 scTCR resulted in a marginal

increase in expression, and the emergence of a well defined monomer peak in size exclusion purification (Figure 3.3), but still yielded low overall levels of expression. Reversing the orientation of the chains, such that the beta chain was located upstream of the linker increased production by greater than an order of magnitude. Expression of DO11 reinforced the necessity of both solubility inducing mutations and β-linker-α orientation, as the orientation alone was insufficient to induce high-level expression.
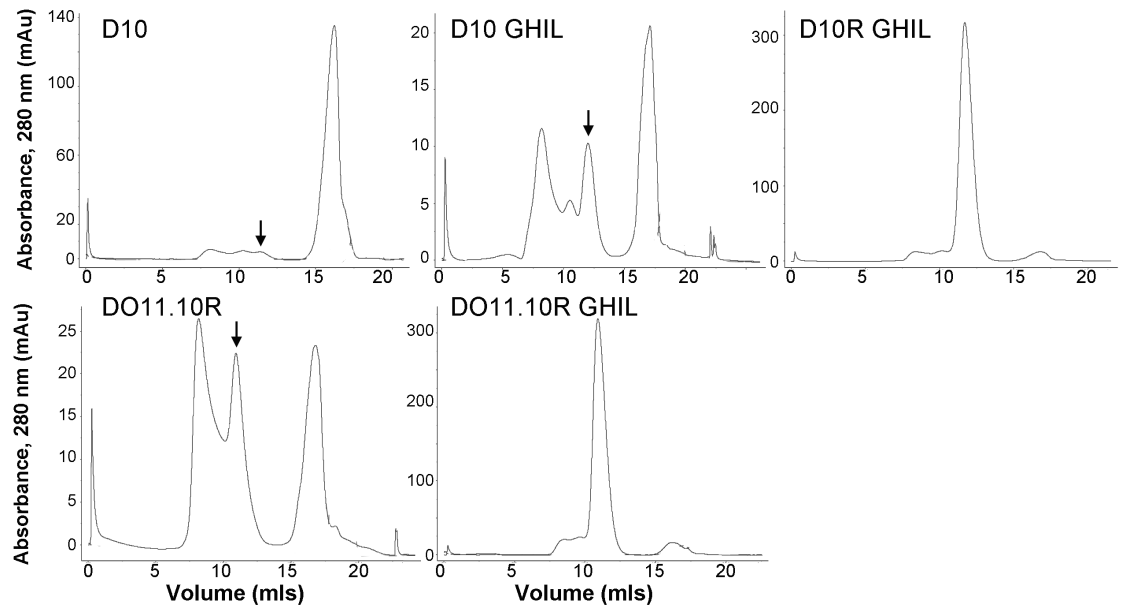
Figure 3.3. Orientation and Solubility-Mutant Effects in Production of DO11 and D10 scTCR.

Protein was purified from 100 mL BL21 *E. coli* culture by osmotic shock and Ni+2 resin prior to column purification. *Top Left*, Chromatogram of original D10 construct with Vα-linker-Vβ orientation and lacking solubility- inducing mutations, induced for four hours. *Top Middle,* D10 with the addition of four solulbility inducing point mutations in TRBV13-2 after four hour induction. *Top Right*, Expression levels following reversal of chain order and 24 hour induction. *Bottom left,* DO11 scTCR in reverse orientation, lacking solubility inducing mutations, 20 hour induction. *Bottom Middle,* DO11 in reverse orientation with solubility inducing mutations. The arrow indicates location of the monomer.

The functionality of DO11 was demonstrated by binding KJ1-26, which recognizes the binding sites of properly paired DO11 alpha and beta chains. DO11 showed a dose dependent response to immobilized KJ1-26, as detected by α−His conjugated to HRP (Figure 3.4). The protein also appears to be stable, as heat treatment at 37 °C for up to 90 hours showed no decrease ELISA response (data not shown).
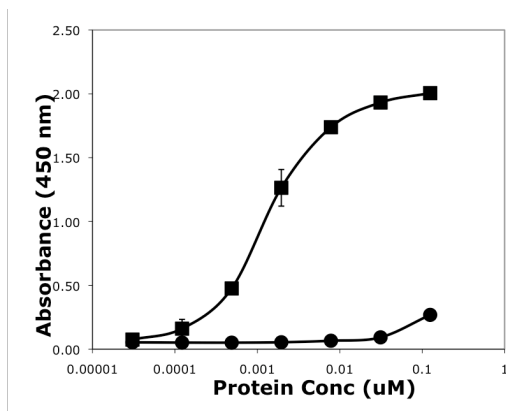
Figure 3.4. DO11 scTCR is Recognized by KJ1-26 α-Idiotype Antibody.

ELISA Plate was coated with 1 μg/mL KJ1-26. DO11 (squares) and D10 (circles) scTCR were added in 3-fold dilution series and washed following 1 hour incubation. Protein binding was detected with α-His HRP.

Based on principles derived for DO11 and D10, all subsequent scTCRs were oriented in this reverse arrangement, and the four described mutants were incorporated into TRBV13-2 containing constructs; the data in Table 3.1 reflects these design principles.

Later, when evaluating the binding of DO11 to its target antigen, we saw fit to examine the solubility-inducing mutants individually. Here we found that reversal of H54Y had no effect on soluble expression, while reversal of either G17E or the pair of I90T, L93S severely reduced expression (data not shown). However, we also find that the expression of DO11 conjugated to M13 pVIII in

phage display (described later) was extremely dependent upon a tyrosine at position 54. In the absence of this mutation, phage titers were extremely low (data not shown). It seems likely that the tyrosine at this position interacts favorably with fusion partners utilized in phage display and the yeast display library from which the mutant was originally isolated.

Despite the relative success of expression for scTCRs containing TRBV13-2, the presence of this gene alone is insufficient to ensure high-level expression, even when properly oriented and mutated. Of particular note, DO11 production was particularly low (less than 5 mg/mL) when expressed from the pMoPac54 vector. pMopac54 is derived from pAK400, but creates a fusion from the c-terminus of the gene to the constant kappa region of light chain human IgG1 (161) In addition, expression was found to be very low in the vector pMoPac10 (161), which produces a protein identical to that of pAK400 (from which it was derived) save for the addition of a "myc" epitope tag at the c-terminus (Kevin Entzminger, personal communication). Soluble DO11 expression was also very low in the previously described pAK500 vector (65), which fuses the c-terminus of the protein to a naturally dimerizing helix domain (data not shown). The sensitivity to c-terminal additions is also seen in the expression of some peptide fusions, described later.

3.3.2.2. scTCR Grafts

Based on the high level expression of several scTCRs, we undertook two efforts to express pairs of TCRs as grafts. Grafting has been widely used as a technique for increasing the stability and expression of antibody fragments (164). Both of our efforts were ultimately abandoned due to poor expression of the resulting constructs. In the first case, we sought to use DO11 as a scaffold for the production of 2B4. TCR functionality is largely dictated by the interaction of six complementarity-determining regions, 3 per chain, with target peptide MHC. Since 2B4 is expressed at extremely low levels, we reasoned that we might

increase its expression by fusing its CDRs to the DO11 framework. DO11's framework was left in tact save for these CDRs, as well as residues outside the designated CDR identified as important for preservation of canonical loop structures by Al-Lazikani et al (165). The expression of the resulting gene was no greater than 2B4 alone (data not shown).

Our second attempt at grafting was less ambitious. We sought to mix and match complementarity-determining regions of 172 and DO11 to enumerate the role of each region in recognition of peptide and MHC recognition. DO11 recognizes ovalbumin 323-339 when presented by I-A^d. This peptide also binds with high affinity to 172's cognate MHC, I-A^u. DO11 and 172 are quite similar. They both incorporate TRBV13-2, meaning their CDR1β and CDR2β regions are identical. In terms of peptide/MHC recognition, therefore, the only differences lie in the respective CDR3s as well as CDR1α and CDR2α. Since CDR1 and CDR2 mediate MHC recognition, while CDR3 is thought to impart peptide specificity. We therefore hypothesized that we could increase the affinity of DO11 for I-A^u by replacing its CDR1α and CDR2α with those found in 172. We synthesized this construct and called it DO72 (Figure 3.5A). We also wondered whether we might impart specificity for I-A^u in conjunction with OVA 323-339 to 172 by replacement of its CDR3α and CDR3β with those of DO11. We synthesized this construct and called it 17DO. It should be noted the prospects of 17DO seemed less attainable, given the propensity of peptides to occupy multiple conformations in I-A^d, and the unique orientation of 172 towards its cognate antigen ((166)).

Since we routinely express both DO11 and 172 at levels in excess of 30 mg/L, the prospect of expressing high levels of the recombined variants of the two seemed feasible. Complementarity determining regions are by definition solvent exposed loops. In addition, the CDR1α and CDR2α loops that differ between the two TCRs share common predicted canonical conformations as predicted by Al-Lazikani et al (165). Nonetheless, expression of both 17DO and DO72 was significantly reduced as compared to their wild type counterparts. In

72

the case of DO72, soluble expression of the monomer was undetectable; for 17DO, it was 3-5 fold reduced (Figure 3.5B).
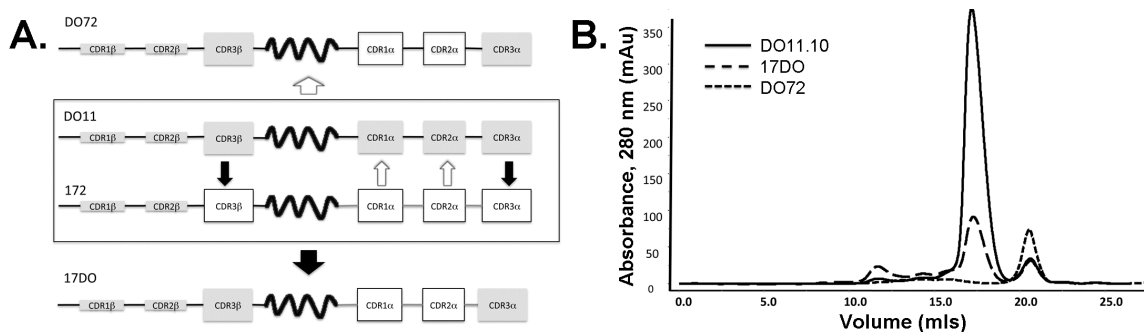
Figure 3.5. Grafts of 172 and DO11 scTCRs

A, **Composition of 17DO, in which CDR3α and β of 172 are replaced with the corresponding amino acids from DO11, and DO72, in which DO11 CDR1α and CDR2α are replaced with corresponding amino acids from 172.** *B,* **Shown are traces resulting from scTCRs purified from 100 mL of culture over an S200 size exclusion column. Expression of DO11 (solid line) is compared with the production of 17DO (dashes), and DO72 (square dots).**

3.3.2.3. Multimerization

To increase the avidity of our scTCRs and render binding to cognate pMHC more sensitive, we sought to create multivalent complexes. Tetramers can be simply created by the natural interaction of four biotins with avidin or one of its derivatives. To facilitate biotinylation of our scTCRs, we cloned two peptide sequences. The first, described as the most efficient substrate for the biotin ligase (167, 168) is referred to here as BSP1 (biotin ligase substrate peptide 1, Table 3.2). The second is commonly used tag also identified by Schatz *et al.*, referred to here as the BSP2. In these tags, the lysine is specifically biotinylated by the enzyme. In addition to these enzymatic sites, we cloned a peptide selected for binding to streptavidin with nanomolar affinity, referred to here as BMP (Biotin mimic peptide, (169)).

Tag	Sequence	Reference
BSP1	GLNDIFEAQKIEWHE	Schatz et al. 1993
BSP2	LHHILDAQKMVNHR	Schatz et al. 1993
BMP	ERCWYVMHWPCNA	Rice et al., 2006

Table 3.2. Peptide Tags for Tetramer Production

Addition of BSP1 to the *N-* terminus significantly increased expression of DO11, while *C-* terminal addition moderately decreased expression of the scTCR (Figure 3.6). *N-* terminal addition of BSP2 proved neutral for protein expression (data not shown). The addition of BMP at either terminus ameliorated expression of the protein monomer (Figure 3.7). This is likely due to the presence of cysteine residues utilized to create the shape of the tag that interacts with the streptavidin binding pocket, as additional cysteines often have the effect of inducing protein aggregation through misfolding.
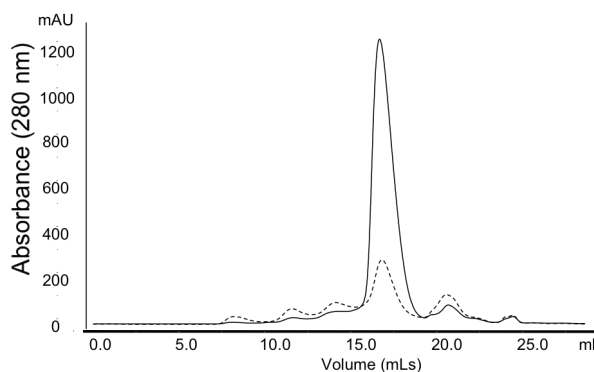

Figure 3.6. Expression of BSP1-linked DO11.

Shown are traces resulting from scTCR fusions purified from 500 mL of culture over an S200 size exclusion column. Short peptide sequences were tethered to the n-terminus (solid lines) or c-terminus (dashed) of single chain DO11, both separated from the TCR on either terminus by a (GGS)$_2$ linker.
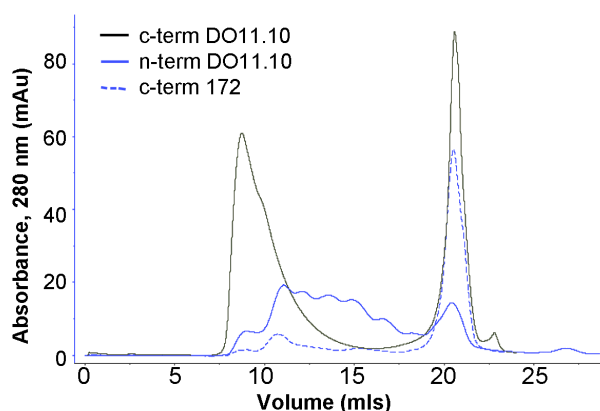
Figure 3.7. Expression of BMP-linked 172 and DO11.

Shown are traces resulting from scTCR fusions purified from 100 mL of culture over an S200 size exclusion column. Streptavidin peptide was fused to the *N-* terminus (blue) or *C-* terminus (black) of DO11, as well as the *C-* terminus of 172 (blue-dashed).

DO11 incorporating *N-* and *C-* terminal addition of BSP1 was enzymatically biotinylated using commercially available biotin ligase and free biotin. Excess biotin was dialyzed away, and proteins were evaluated for biotinylation efficiency using a western blot (Figure 3.8). Biotinylated DO11 was mixed with excess extravidin and run on an SDS-page gel, where DO11 bound to extravidin ran at significantly higher molecular weights. Protein was detected using α-His conjugated to HRP. Both *N-* and *C-* terminal DO11 showed biotinyation efficiencies only approximating 50%. The lower overall detection of *C-* terminal sequence 85 biotinylated DO11 in this blot is likely due to obstruction of the His tag. The formation of DO11 tetramers using biotinylated BSP1 also proved inefficient. Following gradual addition of extravidin, the complex was run on S200 size exclusion column (Figure 3.9). The result appeared to be complexes of mostly TCR dimers in addition to the uncomplexed monomer resulting from unbiotinylated TCR. The low efficiency of biotinylation and subsequent tetramer formation may be due in part to inaccessibility of the sequence tags due to association with the scTCR.

80 kDa ▶

30 kDa ▶

Extravidin: **+**   **+**   **+**
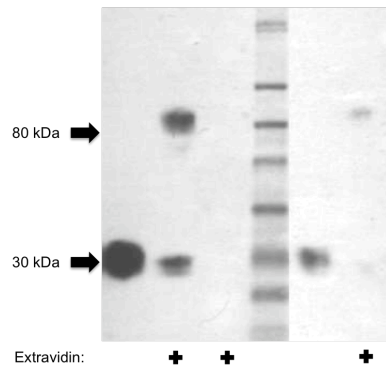
Figure 3.8. Biotinylation Efficiency of BSP-tagged DO11.

5.0 μg of biotinylated DO11 with the BSP1 biotinylation sequence *N-* **(***lanes 1,2***) or** *C-* **(***lanes 5,6***) terminally positioned was separated by a 4-20% gradient gel with (***lanes 2,6***) or without (***lanes 1,5***) the presence of equimolar and detected with** *α*-**His HRP. Extravidin alone was run in lane 4.**
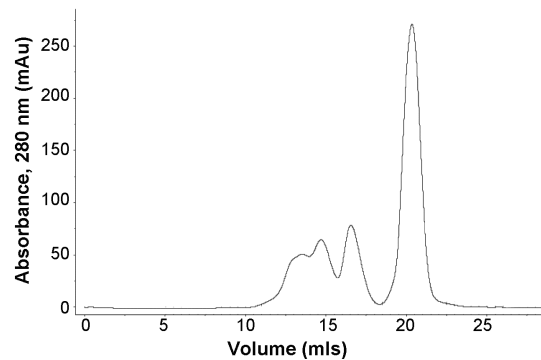
Figure 3.9. DO11 Tetramers.

N-terminally biotinylated DO11 was gradually mixed with extravidin to a 1:4 molar ratio and separated over S200 size exclusion column. Primary peaks from left to right correspond to dimer-extravidin, monomer extravidin, unconjugated T cell receptor, and buffer.

As an alternative to biotin-mediated tetramerization, we pursued dimerization via the use of an epitope tag. The amino acid sequence EYMPME (abbreviated EE) is a truncated version of the larger CEEEEYMPME epitope found in a polyoma virus antigen (170). A monoclonal antibody against this

antigen binds with relatively high affinity for an epitope tag, and the binding of a single antibody to two tag-containing scTCRs would present these in a dimerized form. We chose to place the EE tag at the *N-* terminus of the linker separating the beta and alpha chains (Figure 3.10). It was thought that locating the tag in the middle of the construct would minimize any obstruction associated with terminal placement. Insertion at the *C-* terminus of the beta chain rather than the *N-* terminus of the alpha chain was chosen because of the previously observed insensitivity of TRBV13-2 to the addition of extra peptides. In addition to the addition of a single EE tag, we designed an insert with two tags in succession, separated by an additional Gly_4Ser subunit. This was designed to maximize avidity of the tag towards its antibody, as linear epitope tags at times lack the affinity of other antibody-antigen interactions. Two proximal tags might increase rebinding of the antibody after dissociation, while their proximity to one another should preclude a single antibody from binding to both tags as a sterically unfavorable interaction.
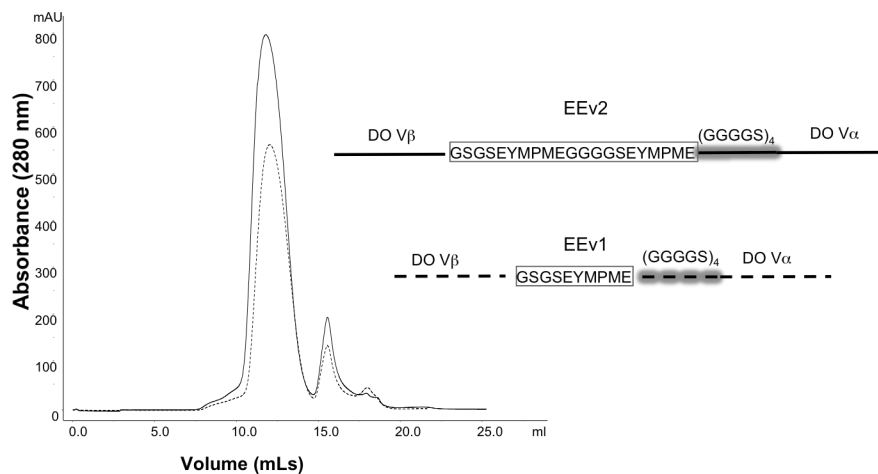
Figure 3.10. Insertion of "EE" Epitope Tags.

Schematics on the right indicate location of EE tag insertion relative to alpha and beta variable region chains and Gly-Ser linker on the TCR. The accompanying chromatogram (left) shows relative production of each scTCR fusion purified from 250 mL of culture over an S200 size exclusion column for the EEv1 (dashed) and EEv2 variants.

77

The insertion of both forms of the EE tag was well tolerated by DO11, and both were produced at high levels (>30 mg/L, Figure 3.10). The efficacy of this tag in producing functional dimers was evaluated using an ELISA. Plates were coated with KJ1-26. Dimers were assembled by gradual addition of monoclonal α−EE to a molar ration of 1:2, incubated, and detected by the addition of α-IgG1 conjugated with HRP. Both versions of EE-tagged dimers were functional, with the EE-doublet showing slightly greater EC_{50} (Figure 3.11).
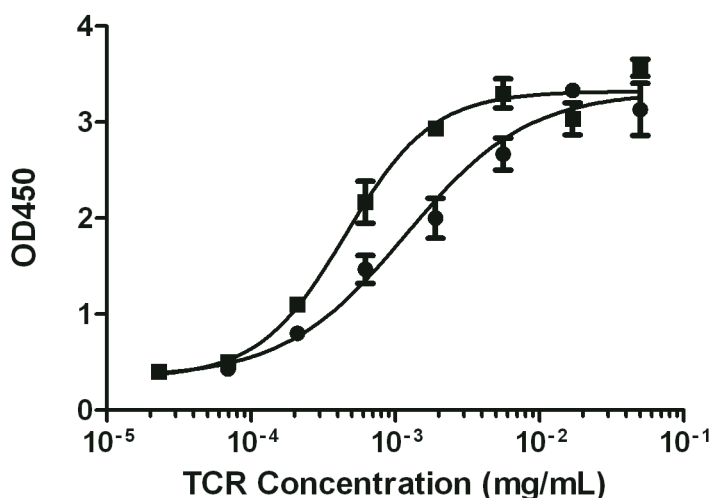
Figure 3.11. Activity of EE-Tagged DO11 Dimers.

ELISA wells were coated with 1 μg/mL KJ1-26. DO11 EEv1 (circles) and DO11 EEv2 (squares) dimerized with 1:2 molar α-EE were added in 3-fold dilution series and washed following 1 hour incubation. Protein binding was detected with α-IgG1 HRP.

3.3.2.4. IL10 Peptide Mimic

One potential use for scTCRs is as a targeting mechanism for specific peptide antigens displayed by the immune system. Interleukin 10 (IL-10) is a cytokine that is broadly anti-inflammatory, and can shift immune responses from a cell mediated inflammatory response to a humoral response (171) Given that overactive inflammatory response has been implicated in diseases such as

multiple sclerosis, psoriasis, rheumatoid arthritis, and inflammatory bowel disease, the use of IL10 as a therapeutic has been variously explored. Recently, a nonapeptide mimic with complete homology to the *c-* terminal portion of human IL10 has been shown to possess many of the functional properties of the complete cytokine (158-160). We decided to test the compatibility of this short peptide with our scTCRs.

The IT9302 peptide mimic (AYMTMKIRN) was cloned to express as a fusion protein in several ways. First, we directly linked to the *n-* terminus of DO11 and 172. Second, we coded the sequence directly into pAK400 between the SfiI restriction site used for scTCR and the 6-His epitope tag. Expression was of both the *n-* terminal (Figure 3.12A) and *c-* terminal fusions (Figure 3.12B) were reduced compared to unmodified DO11 and 172 constructs. As an additional test, we created a version of the vector in which the peptide was expressed after the His Tag. Expression of this construct was increased relative to unmodified 172 (Figure 3.12C).
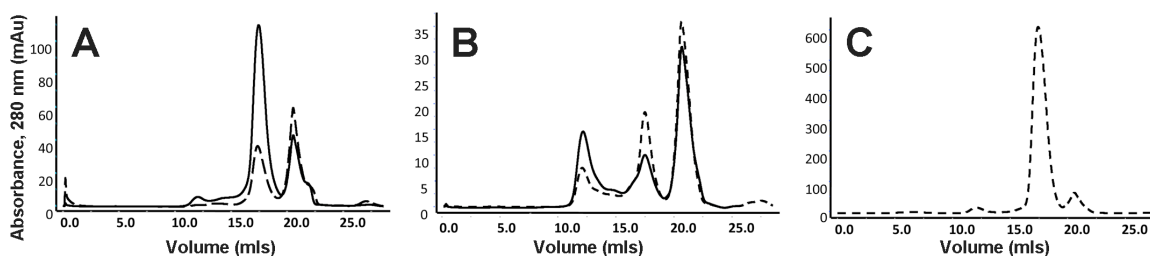
Figure 3.12. Fusion of IL10 Peptide Mimic.

DO11 (*solid lines*) and 172 (*dashed lines*) were expressed with the IL10 peptide mimic IT9302 tethered to the scTCR by a (Gly$_2$Ser)$_2$ linker at the n-terminus (*A*) or c-terminus (*B*) prior to the His tag. For 172, the peptide was also tethered after the His tag (*C*). Each curve produced from S200 size exclusion chromatography of Ni^{2+}isolated protein from 100 mL of *E. coli* culture.

3.3.3. Phage Display

Given their relatively low affinities for target antigens, affinity maturation was pursued to increase the utility of scTCRs as reagents or therapeutics. Phage display was previously discussed as a means for affinity maturation (Section 1.2.3). In the following section, we tested the applicability of current phage display systems for use with DO11 and 172.

3.3.3.1. pIII Display of 172

Initially we tested the display of 172.10 in the pAK100 and pMoPac24 display vectors that fuse the displayed protein to a truncated version of M13 protein III (pIII). Particles were produced through induction of gene expression with IPTG and infection of "helper phage" containing complete viral genomes (65). Titres of approximately 10^{11} phage/ mL were produced and concentrated using 20% PEG-2.5M NaCl precipitation. Phage were resuspended in 50% glycerol in PBS and stored at -20° C.

Expression of scTCR on the particles was demonstrated via ELISA. Both pAK100 and pMoPac24 code for a "c-myc" amino acid tag between the protein and its fusion partner. Immobilized α−myc was used to pull down phage particles, which were then detected using α−pVIII, which binds to the major protein of the phage coat. The 172 scTCR showed similar levels of display to the control scFv 14B7, as evidenced by a dose response to number of phage particles present (Figure 3.13A). 172 and 14B7 also produced similar phage titers (data not shown). 172 was more effectively displayed using the pMoPac24 vector, which coexpresses the skp chaperone protein (Figure 3.13B).
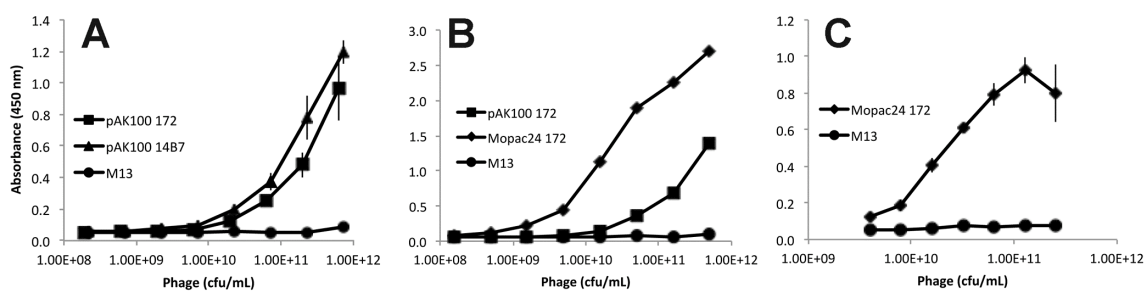
Figure 3.13. pIII Display of 172.

172 scTCR was displayed as an M13 pIII fusion and panned against immobilized 1 mg/mL α-myc (*A*, *B*) or 0.4 μM MBP1-11 covalently linked to I-A^u (*C*). Binding was detected by α-pVIII HRP.

To demonstrate functionality of the displayed scTCR, particles were panned against immobilized MBP ^{Ac}1-11 covalently linked to I-A^u. The interaction between MHC-peptide and TCR (K_d ~5 uM, (163)) is of lower affinity than that of the antibody/tag, yet we still observe dose-dependent interaction (Figure 3.13C).

3.3.3.2. Protein VIII Display

For some lower affinity scTCRs, display on pIII may be insufficient to allow for target binding, given the absence of any avidity effect. pVIII, on the other hand, comprises the bulk of the phage coat, and on the order of 1000 copies are present per phage. Based on this successful display of 172 in pMoPac24, we opted to modify this vector to fuse incorporated proteins to full-length pVIII.

Despite high copy numbers of pVIII present on each phage particle, previous efforts at displaying larger fusion proteins has produced mixed results (172, 173). This difficulty is likely due to the close-packed nature of this protein, and the necessary close association of a displayed protein with its fusion partner. To improve display via pVIII, Sidhu *et al.* constructed five libraries of protein VIII variants and selected for the display of human growth hormone (174). Each library consisted of random mutants spanning ten amino acids, one for each of

81

the five segments of this 50 amino acid protein. From these, he selected several optimized variants for each of the first three n-terminal most segments. In a separate publication (175), the utility of each individual mutation was evaluated. Given the possibility that the increased display demonstrated by these mutants was based proximity to and specific interactions with the HGH used in selection, we opted to synthesize all three of the described pVIII variants and test their efficacy in the display of DO11 (Table 3.3).

Region 1									
Amino Acid 1	2	3	4	5	6	7	8	9	10
pVIII WT A	E	G	D	D	P	A	K	A	A
pVIII o1 A	K			E	K	F		R	D

Region 2									
Amino Acid 11	12	13	14	15	16	17	18	19	20
pVIII WT F	N	S	L	Q	A	S	A	T	E
pVIII o2 Y	E	A			E	D	I		

Region 3									
Amino Acid 21	22	23	24	25	26	27	28	29	30
pVIII WT Y	I	G	Y	A	W	A	M	V	V
pVIII o3 L	F	F	L	L	G	T	V	H	L

Region 4									
Amino Acid 31	32	33	34	35	36	37	38	39	40
pVIII WT V	I	V	G	A	T	I	G	I	K

Region 5									
Amino Acid 41	42	43	44	45	46	47	48	49	50
pVIII WT L	F	K	K	F	T	S	K	A	S

Table 3.3. Variants of pVIII for phage display.

Three variants were cloned into pMoPac24 for fusion based phage display based on the work of Sidhu *et al.* (174). Mutations in each variant were limited to a single 10 amino acid segment.

Wild-type protein VIII and the three optimized variants were synthesized via PCR amplification of oligonucleotides with flanking Asc I and Hind III restriction sites. These fragments were ligated into pMoPac24 in place of the pIII fusion to produce pMo24p8, pMo24p8o1, pMo24p8o1, and pMo24p8o3. We then cloned DO11 and control scFv 14B7 were cloned into each of the variants and assessed for display of the fusion protein.

82

Vectors were transformed into XL1-B, and phage particles were produced by infection of M13KO7 and concurrent induction with IPTG, followed by expression for up to 24 hours. Phage titers were evaluated as colony forming units, and varied greatly among the variants. For 14B7, the pVIIIo1 fusion produced the highest titers, followed by pVIIIo2 and pVIII WT, with pVIIIo3 showing very low levels of production. For DO11, on the other hand, titers of the wild type and pVIIIo2 were roughly equal, with significantly lower titers shown for pVIIIo1 and pVIIIo3. As a whole, titers ranged from nearly equal magnitude to pIII display to over four orders of magnitude lower Figure (3.14).
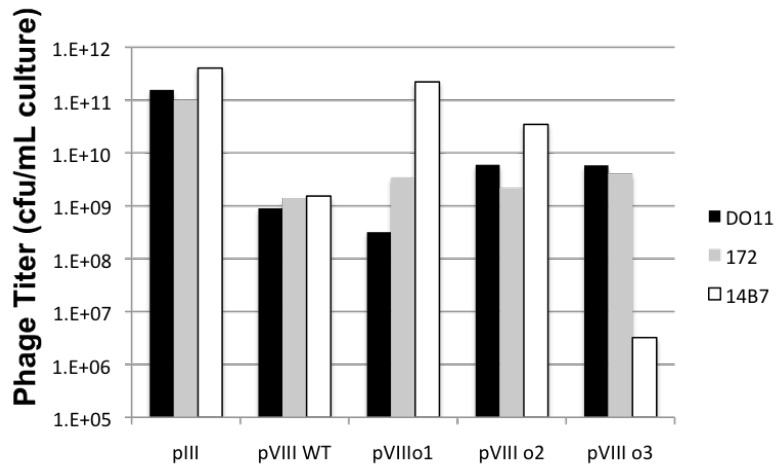
Figure 3.14. pVIII-associated phage titers are specific to the displayed protein.

Next, particles were evaluated for the display of the fusion protein (Figure 3.15). Phage was panned against immobilized binding target, KJ1-26 for the case of DO11, and anthrax protective antigen (PA) for 14B7. Binding was detected by α-pVIII HRP. Both proteins showed slightly lower display via pVIIIo1 when compared to the wild type fusion. For DO11, both pVIIIo2 and pVIIIo3 showed significantly improved display; for 14B7, pVIIIo2 display was similar to that of the wild type, while pVIIIo3 titers were too low to judge.
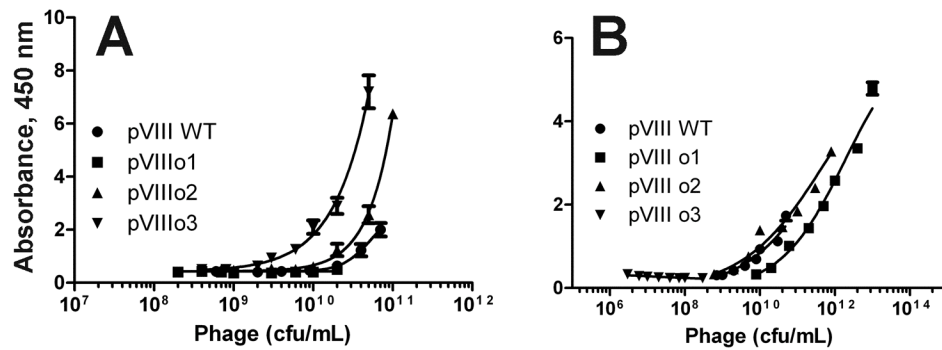
Figure 3.15. pVIII Display of DO11 and 14B7.

DO11 scTCR (*A*) and 14B7 scFv (*B*) were was displayed as on M13 fused to pVIII (*circles*) or an optimized variant and panned against immobilized 1 μg/mL KJ1-26 or anthrax protective antigen. Binding was detected by α-pVIII HRP.

To compare the effective avidity of pVIII and pIII display, phage particles were used to block the interaction of DO11 cells with A20 antigen- presenting cells (Figure 3.16). Phage particles were added to 96-well plates in a dose-dependent manner. Stimulatory peptide (20 μM) was added to 15,000 A20 lymphoma B cells presenting I-A^d and 15,000 DO11 Hybridoma cells, and allowed to incubate for 24 hours. Phage displayed variants pVIIIo2 and pVIIIo3 were found to inhibit the production of interleukin 2 by DO11 cells at titers of approximately two orders of magnitude lower concentration than pIII displayed phage. Particles displaying pIII, on the other hand, were similar in efficacy of IL2 inhibition to M13 phage particles alone.
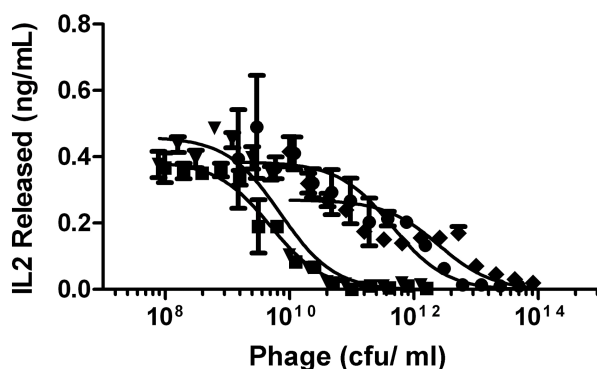
Figure 3.16. Inhibition of IL2 Release by pVIII-displayed DO11.

Phage particles alone (*diamonds*) or displaying DO11 scTCR on pIII (*circles*), or pVIII variants o2 (*squares*) or o3 (*inverted triangles*) were incubated in the presence 20 μM 323-339 peptide along with 15,000 DO11 T cell hybridoma cells and 15,000 A20 B Cell lymphoma cells. Release of IL2 into the supernatant was quantified by matched pair ELISA.

3.3.4. Affinity Maturation of DO11.

3.3.4.1. Selection of Phage Display System

The first step in the maturation of DO11 was in choosing an antigen to pan against. Typically, an isolated peptide covalently linked to an MHC would be appropriate, covalently linked ova 323-339 I-A^d has proven ineffective at staining DO11 T hybridoma cells, for reasons discussed in Chapter 2. As a result, maturation against the α-idiotype antibody KJ1-26 was used. Due to the higher affinity of this interaction relative to the native TCR-pMHC, pIII-based display was first attempted.

As a demonstration of the feasibility we initially showed enrichment of particles with DO11 displayed from severe dilution within 172-displayed particles. DO11 particles were diluted at a rate of 1:1000 or 1:1000 with pIII-displayed 172 (Figure 3.17). The combined mix of particles was panned against KJ1-26 or a non-specific antibody, eluted with a solution of low pH glycine, neutralized with a

high pH Tris buffer, reinfected into ER2738 cells and amplified overnight prior to another round of panning. For DO11 phage alone, between 10^4 and 10^5 colony forming units were eluted per between 10^9 and 10^{11} cfu inputted to each round of panning. For DO11 diluted one to a million in 172 phage, on the order of 100 particles were selected after a single round of panning, increasing to 10,000 after four successive rounds of selection and amplification, indicating that pIII displayed DO11 is effectively selected for in panning against KJ1-26.
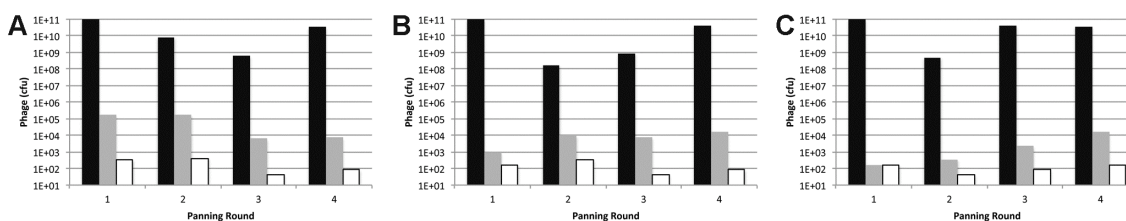
Figure 3.17. Enrichment of DO11 Phage Panned Against KJ1-26.

DO11 scTCR displayed as a pIII fusion alone (*A*) or diluted 1:1,000 (*B*) or 1:1,000,000 (*C*) in pIII- displayed 172 particles was panned against immobilized KJ1-26 (1 µg/mL). *Black* columns show total phage inputted to each round of panning, *gray* columns show total phage eluted from following KJ1-26 binding and *white* columns show total phage eluted from wells coated with control antibody.

3.3.4.2. Selection of Affinity Matured Mutants.

Site directed mutagenesis was used to randomize the CDR3α region of DO11. Prior to randomization, three consecutive "stop" codons were initially introduced into the wild type DO11 using full amplification of the vector. These stop codons prevent wild type DO11 from outcompeting variants, as each individual variant is present in low quantities prior to any selection. Single stranded DNA was then isolated from phage carrying the stop codon-laden DO11 template. Libraries were created by the slow annealing of oligonucleotides containing random codons to this template followed by extension of the primer around the plasmid at room temperature and ligation. Libraries were then

transformed via electroporation into XL1B cells. Cells were then infected with helper phage, induced to express protein with 1 mM IPTG and grown overnight to produce phage libraries.

Two separate libraries were created, each introducing six random codons into CDR3α. The first preserved the 10 amino acid length of the wild type CDR, randomizing from position 107 through 115 (standard IMGT numbering includes amino acids not present in some TCRs, Table 3.4), preserving the reasonably well conserved alanine at position 106 and leucine at 116. The second library randomized positions 106 and 116, but created a shorter overall CDR3α of 8 amino acids. "NNS" codons were used for randomization, where N can be any or the four possible nucleotide bases and S is restricted to guanine or cytosine, to limit the number of stop codons within the library. Each library contained approximately 10^6 variants and was panned in five successive rounds against immobilized KJ1-26.

During the course of panning, the 8 amino acid- length library was eventually overtaken by contaminating variants from the 10 amino acid-length library, likely indicating a lower inherent affinity in the shorter library. Both libraries were also contaminated by wild type DO11, which was unable to outcompete several members of the 10 amino acid library. Selected DO11 variants are presented in Table 3.4.

	104	105	106	107	108	109	113	114	115	116	117
DO11	C	A	A	S	P	N	Y	N	V	L	Y
1721	C	A	D	D	Q			S	N	R	Y
17210	C	A	A	G	H			R	C	N	Y
1328	C	A	A	S	S	F	T	P	R	L	Y
235	C	A	A	S	Y	A	Y	H	Q	L	Y
176	C	A	A	S	T	N	W	H	N	L	Y
174	C	A	A	A	L	N	Y	H	S	L	Y
2343	C	A	A	V	P	N	Y	H	L	L	Y
2146	C	A	A	T	P	W	L	S	R	L	Y
1746	C	A	A	P	R	T	A	F	N	L	Y

Table 3.4. Selected Variants from DO11 CDR3α Library.

Gray residues were unmodified by random codons. Positional numbering is according to IMGT. Bolded variants were cloned into pAK400 for further analysis

Three mutants were chosen to for further comparison to DO11. Each was subcloned into pAK400, expressed, and purified from BL21 cell culture. Yields were comparable to DO11 (data not shown). Each was then tested in ELISA against KJ1-26, with detection by α-His HRP, with two of the three mutants showed significantly higher EC_{50} (Figure 3.18).
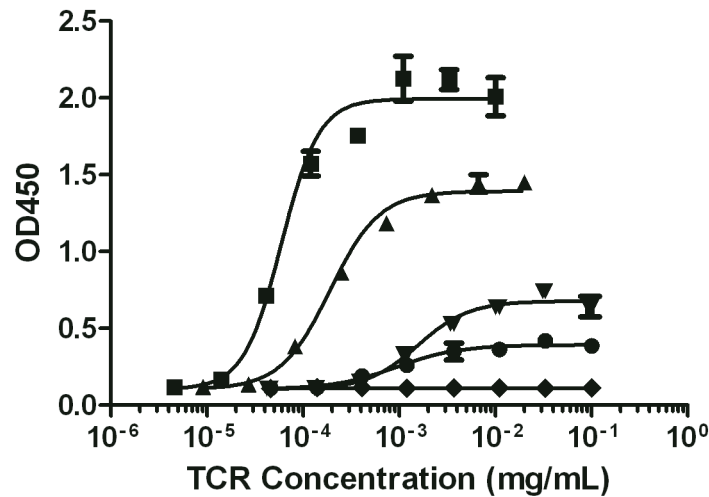
Figure 3.18. Activity of Affinity-Matured DO11 Variants.

Affinity matured variants of DO11 were cloned into pAK400 and expressed as soluble proteins. Following purification, protein doses were incubated with immobilized KJ1-26 coated at 1 μg/mL in 96 well plates. Binding of DO11 (circles) 235 (inverted triangles) 174 (triangles) and 176 (squares) was detected with α-His HRP. Diamonds indicate the average of all samples bound to an irrelevant protein.

Variants 176 and 235 were further measured for affinity against KJ1-26 with BIACore. KJ1-26 was immobilized and monomer DO11 and variants were flowed across the surface with the antibody, while changes in resonance were monitored for the determination of on- and off-rates. These results paralleled the capture ELISA, with 176 showing approximately two orders of magnitude higher affinity than DO11 (Table 3.5). The increase in affinity was primarily due to slower off rate (Figure 3.19). Variant 235 showed approximately equal affinity to DO11.

89
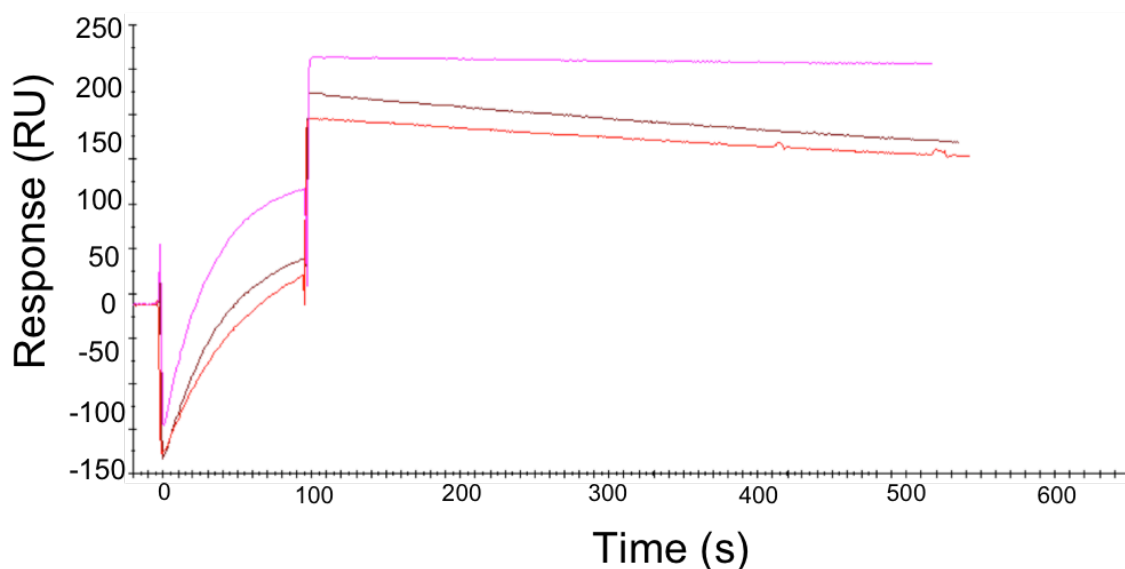
Figure 3.19. Determination of DO11 Variant Affinity with BIACore.

Purified DO11 scTCR and variants were assessed for affinity to KJ1-26 with BIACore 2000. Shown are representative curves of 176 *(pink)*, 235 *(brown)*, and DO11 *(red)* at 200nM flowed at 35 μL/min over 600 RU KJ1-26 immobilized to chip.

Species	Ka (1/Ms)	Kd (1/s)	KD (M)
DO11.10	2E+05	3E-03	**4E-08**
235	4E+04	6E-04	**2E-08**
176	1E+05	3E-05	**2E-10**

Table 3.5. Affinity of DO11 Variants.

Kinetic data were fitted for association and dissocation rate constants from duplicate, two-fold dose response sereies descending from 400 nM using BIAEvaluation software.

Based on the higher affinity of 176 against KJ1-26, this variant was tested for its ability to stain A20 antigen-presenting cells. Two consecutive EE tags were inserted prior to the linker, as described in Section 3.3.2.3, and the protein was expressed in BL21 cells. Dimers were formed from purified monomer by the stepwise addition of monoclonal α−EE to a 1:2 molar ratio of antibody to scTCR.

Fluorescent tetramers were further formed from dimers by the gradual addition of allophycocyanin (APC)-conjugated monoclonal α-IgG1 (Figure 3.20).
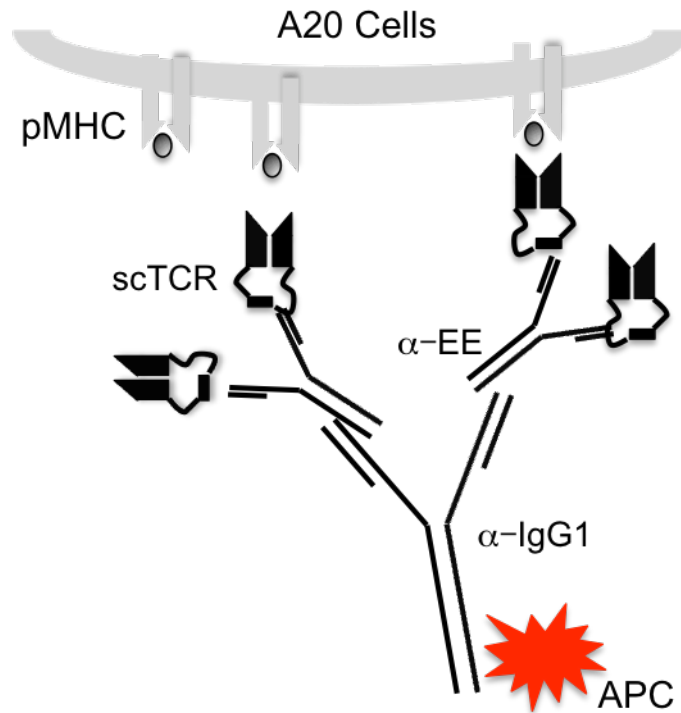
Figure 3.20. Representation of Assembled, EE Tag-based Tetramers.

A20 cells (250,000) were then incubated with 50 μM peptide along with 0.5 μg DO11 or 176 in tetramer form. EE-tagged, tetramerized maltose binding protein was used in equimolar amounts as a control. Results indicate that 176 stains cells more efficiently than DO11 (Figure 3.21B). Somewhat surprisingly, given the antibody substitute used for affinity maturation, 176 also exhibits peptide specificity (Figure 3.21C) in staining APCs, which is not seen for DO11 (data not shown).
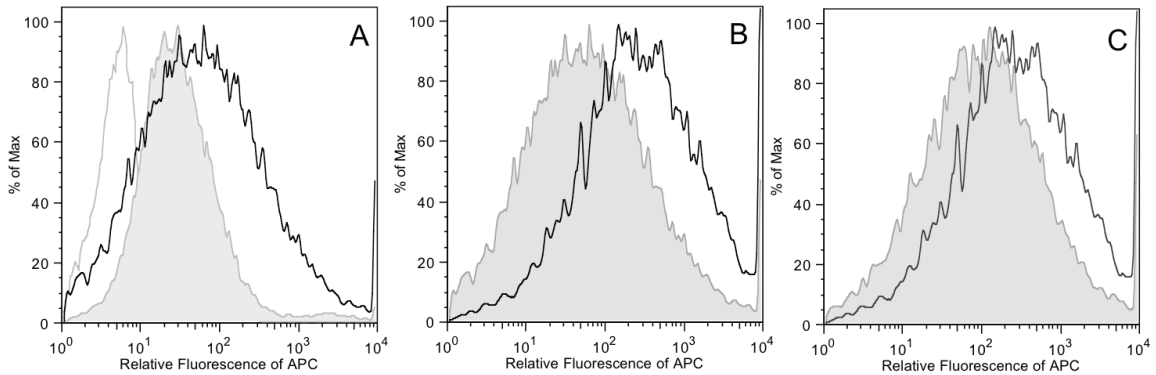
Figure 3.21. Fluorescent Staining of A20 Cells with DO11 and 176.

Purified scTCR (0.5 mg) or equimolar EE-tagged Maltose binding protein bound to α-EE and α-IgG1 was used to stain A20 lymphoma cells in the presence of 50 mM ova 323-339 peptide. *A*: DO11 (*black*) is compared with unstained control (*gray*) and maltose binding protein (*gray, filled*). *B*: 176 (*black*) is compared with DO11 (*gray, filled*). *C*: 176 is shown with (*black*) and without (*gray, filled*) the addition of ova peptide.

3.4. Discussion

3.4.1. scTCR Expression

The observed stability of TRBV13-2 containing scTCRs seems not to be unique among variable regions, as a similar effect has been demonstrated for the human TRAV-12 family (176). Neither, it should be noted, is it universal, as solubility-inducing mutations and chain order are important, and various peptide tags significantly decreased expression.

The preference for "β-first" scTCR orientation could result from one of several effects. In the case where alpha and beta chains exhibit significantly different thermodynamic stabilities, the beta chain may drive proper folding of the less stable alpha chain. Stable bacterial proteins such maltose binding protein are commonly fused with difficult to express partners with great success. It may be that when the beta chain is translated first, it can fold rapidly and provide assistance for it's alpha binding partner during folding. A more likely explanation,

in our opinion, can be gleaned from data gathered in our various scTCR fusion constructs. In the standard construct expressed by pAK400, the n-terminus of the initially expressed variable region (α or β) is fused to a pelB leader sequence, which aids in transport of the protein to the periplasm. Likewise, the *c*- terminus is fused to a linker plus the *N*- terminus of the subsequent chain. Given that we observe that α chain of DO11 is relatively intolerant to peptide or fusion protein additions at its at *c*- terminus, that preference probably applies for effective "linker-β" *c*- terminal tag necessitated by the forward conformation. Likewise, the β chain tolerance for *N*- terminal fusions likely carries over to its tolerance of the pelB leader sequence.

The lack of expression of attempted grafts of DO11 (with 2B4 and 172) could be further explored using DNA shuffling or random mutagenesis for insight into previously unidentified amino acid positions important in genetically encoded CDR1 and CDR2 presentation. Notably, the greater expression of 17DO as compared to DO72 (Figure 3.5B) is perhaps reflective of the inherent tolerance of TCRs towards various CDR3 compositions.

3.4.2. Affinity Maturation

The selected variants from the DO11 library showed moderate conservation at for four of the six randomized codons: 107S, 109N, and 113Y (Table 3.5). Of the seven selected variants from the 10 amino acid-length library, at least 3 incorporate a wild-type residue at these positions. A non-native histidine at position 114 was selected in four of the seven unique variants. Further recombination of the selected mutants may yield greater affinity enhancement.

To our knowledge, the variant 176 marks the only the second reported affinity variant of TCR to a Class II peptide MHC. The use of KJ1-26 as a binding partner facilitated the panning, but it is unclear whether other anti-idiotypes would be similarly well suited to serve as proxies for native binding target. Given the

93

differences in the native structure of antibodies and peptide-MHC constructs, it is difficult to imagine the former mimicking the latter with any notable frequency. It is possible, however, that using a small scTCR for mouse immunization or panning against a naïve library would yield such variants with greater frequency.

CHAPTER 4

PRIMER SET FOR THE AMPLIFICATION OF MOUSE TCR CDNA

4.1. Introduction

The description and characterization of a primer set is shown here for the amplification of murine cDNA for use in the high throughput sequencing of the TCR repertoire. Sequencing with the next generation 454 pyrosequencing-based technology is currently being performed with the goal of evaluating changes in TCR usage associated with ovalbumin immunization. The resulting data is being processed, so the current description is limited to a feasibility study in which TCRs were identified from two previously described "Type B" cell lines followed by a small-scale preliminary 454 sequencing run.

High throughput sequencing has the potential to delineate the relationship between T cell receptor sequence and the function of host T cells. Receptors have been suspected of influencing T cell development towards specific subsets based on their inherent binding affinities; high affinity receptors are thought to play a role in the development of regulatory cells (17, 19) and memory cells (177). Recent deep sequencing work in humans, however, has suggested that T cell differentiation is stochastic, as identical receptors are found throughout various subsets (178). In addition, the preferential usage of certain genes has long been noted in the certain immune responses, for instance, usage of TRBV13 in murine models of autoimmune disease (179). Deep sequencing of selected subsets has the potential to resolve these long-held observations as well as to provide a wealth of information about the development of various disease states.

4.2. Materials and Methods

4.2.1. Primer Set Design

Oligonucleotides were designed for the amplification of mouse TCR α and β cDNA based on reference sequences of each gene and listed allele from IMGT (180). The constant region of each chain was targeted with a single α-sense oligonucleotide (α: 5'- GTGCTGTCCTGAGACCGAGG- 3'; β: 5'- GCCTTTTGTTTGTTTGCAATCTCTGC- 3'), which anneal 12 residues into each constant region. Individual sense strand primers were designed to be complementary to catalogued leader sequences (Table 4.1, 4.2) according to principles described below. Primers recognizing pseudogenes (gene fragments with stop codons or frameshifts such that an in tact TCR is not produced) were not included were omitted from consideration in design as not relevant to the in vivo repertoire. Genes from the additional duplication event seen in C57/BL6 mice (marked "N" by IMGT) were also omitted from consideration, as were delta genes not shared with the alpha repertoire. (Figure 4.1)
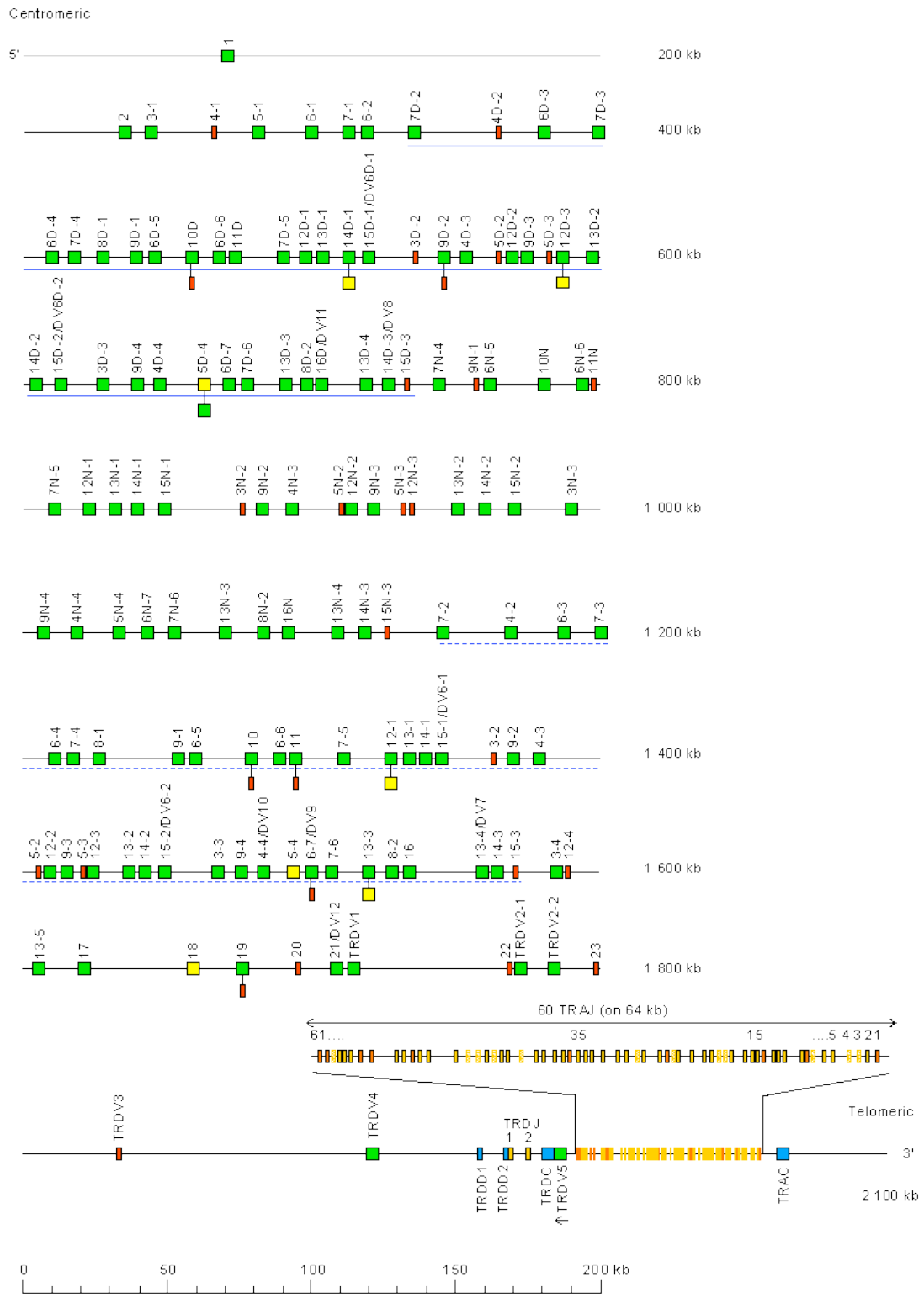
Figure 4.1. Mouse TCRα Locus

97

A representation of the mouse TCRa locus from chromosome 14. Green indicates a functional gene, yellow an open reading frame, and red a pseudogene. Genes denoted with an "N" appear only in mice of the C57/BL6 strain. (Source: (180))

4.2.2. 4C4 and 3B6

Two murine T cell lines, 4C4 and 3B6 reactive to a human myelin basic protein peptide were obtained from Thomas Forsthuber (University of Texas-San Antonio) for the purpose of extracting TCR Sequences.

4.2.3. Immunization and Cell Sorting

Four 25-27 week old BALB/c mice were each immunized with 100 μg of ovalbumin or PBS in CFA in two subcutaneous injections behind the shoulder blades. After 14 days, spleens and lymph nodes were harvested. Lymph nodes were immediately transferred to RNA later (Invitrogen) and stored at -80 C prior to total RNA isolation. Splenic tissue was disrupted with the plunger of a syringe and passed through a cell strainer for subsequent isolation of T cell subsets. Following red blood cell lysis, cells were separated by first isolating $CD4^+$, then $CD8^+$ cells using magnetic Dynabeads (Invitrogen #114.45D, 114.47D) by incubation for 20 minutes at 4 °C with gentle rotation. Beads were magnetically separated from supernatant and stored in TRIzol reagent (Ambion).

4.2.4. TCR Amplification

Total RNA was first purified from isolated cell populations. Cells were lysed in TRIzol (Ambion) through homogenization (OMNI GLH), repeated drawing through 23-gauge needle and syringe, or extensive vortexing (for bead-bound populations). Following the addition of 1- bromo 3- chloropropane and centrifugation, RNA was isolated from the aqueous phase. RNA was further purified through by ethanol precipitation on a column using QIAgen's RNEasy kit. DNA contamination in purified samples was degraded using DNA-free treatment kit (Ambion). Samples were stored

cDNA was obtained through reverse transcription using SuperScript RT II (Invitrogen) extension of oligo(dT) primers for 50 minutes at 42 °C. T cell receptors were amplified from cDNA through amplification of with 25-30 cycles of PCR using the primer set and Accuprime pfx polymerase (Invitrogen). Resulting DNA was separated on a 1.5% agarose gel and purified using QIAgen's QIAquick gel extraction kit. Purified DNA was cloned into Zero Blunt TOPO kit or submitted to the Genome Sequencing and Analysis Facility at the University of Texas for adaptor ligation and subsequent high throughput sequencing using a Roche/454 FLX Titanium Series Instrument.

4.2.5. Data Analysis

Preliminary high throughput sequencing data was submitted to VQuest at IMGT.org (181) using a script developed by Zach Frye (The University of Texas at Austin).

4.3. Results and Discussion

4.3.1. Primer Set Design and Validation

Several principles guided the design of TCR primers. A minimum number of overall primers were desired. Degeneracy was introduced into primers based on available sequence data when necessary, but limited to four or fewer "versions" of each primer (e.g. related primers in which one position is allowed to randomize to all four possible nucleotides). In order minimize the number of primers as well as the degree of degeneracy, the length of each primer was selected such that the melting temperature of each primer was approximately 55 °C. Primers were annealed to leader sequences with the exception of TRBV13, which targeted a conserved region at the beginning of the framework.

	Name	Sequence	Targets (TRAVX)
1	TRAV10	ACATCCCTTCACACTGTATTCC	10; 10D
2	TRAV11	GCCTKAGTGCCTGCTGG	11; 11D
3	TRAV12	SCTCAGTTCTYGTGCTCCTC	12- 1:3; 12D- 1:3
4	TRAV13	TCTCTGYTGGGGCTYCTG	13- 1:5; 13D- 1:4
5	TRAV14	GACAMGATCCTGACAGCAWC	14- 1:3; 14D- 1:3
6	TRAV15	ATGCCTCCTCASAGCCTG	15- 1:2; 15D- 1:2
7	TRAV16	CTGATTCTAAGCCTGYTGGG	16; 16D
8	TRAV17	AGTGACCATTCTGCTGCTC	17
9	TRAV18	ATGCTCCTGAAACTCTCTGTG	18
10	TRAV19	GCCTTGCTGTTGGTTCTG	19
11	TRAV1	GGTTTGTTCTCTATCTCTTCCTG	1
12	TRAV21	GATGTGTGAGTGGAATTGCC	21
13	TRAV2	ATGAAGCAGGTGGCAAAAG	2
14	TRAV3	GACCTTTGTTSCTGTGCTTC	3- 1,3,4; 3D-3
15	TRAV4	GCTGTGCTGGGGATTC	4- 3,4; 4D- 3,4
16	TRAV4-2	CTGCTGTTGGTGCCGCT	4-2
17	TRAV5-1	CTTCTGGCTACAGATGGACT	5- 1
18	TRAV5x-4	ATTATTCATGTTTCTATGGCTGC	5- 4; 5D- 4
19	TRAV6-1to3, 6d-3	GCTTTAGTGACTGTGATGCTG	6- 1:3; 6D- 3
20	TRAV6x-4	GTTTTAGTRACTGYGATGCTGCTG	6- 4; 6D- 4
21	TRAV6x-6	CTYTTCTCCAGGCTTCGTG	6- 6; 6D- 6
22	TRAV6x-7	ATGAYTGTGATGCTCCTCAT	6- 7; 6D- 7
23	TRAV7to7D5	CTAGTGGTCCTGTGGCTY	7- 1:5; 7D- 2:5
24	TRAV7x-6	CTAGTGTTCCTCTGGCTTC	7- 6; 7D- 6
25	TRAV 8x-1	GGTTGTTGATGGTGTCACTG	8- 1; 8D- 1
26	TRAV 8x-2	GGAATATCTTTGGTGACTCTATGG	8 -2; 8D-2
27	TRAV9x-1	TCGTTCCTCGGGATACATTTC	9- 1; 9D- 1
28	TRAV9x-2to4	CAGTSCTGGGGATACACT	9- 2:4; 9D- 2:4

Table 4.1. Primers used in the amplification of TCRα genes

For TRAV, a total 28 primers were designed (excluding multiples of degenerate versions) to target 82 genes. The mouse alpha region features a duplication event (and a triplication event in the C57/BL6 strain), and genes are grouped according to families (Figure 4.1, (182)). Several families contained diversity necessitating multiple primers for amplification of a single family. The average melting temp was 55 °C, ranging from 52-59 °C. For TRBV, 19 primers were designed to amplify 22 genes. Average melting temp was 56 °C with a range of 54-57 °C.

	Name	Sequence	TRBV
1	TRBV1	TTCTGTGCCTCTGTGTACTCAT	1
2	TRBV2	GGCTCCATTTTCCTCAGTTGC	2
3	TRBV3	CTGGCTTCTAGGTTGGATAATTTTTAG	3
4	TRBV4	GCTGTAGGCTCCTAAGCTGT	4
5	TRBV5	GCAGGCTTCTCCTCTATGTTTC	5
6	TRBV12	CTGCTATCTTGGGTTRCTSTCTTTC	12- 1:2
7	TRBV13	RCAAGGTGRCAGTAACAGGAG	13- 1:3
8	TRBV14	CTTGGCTGGGCAGTGTTC	14
9	TRBV15	AGACCCTCTGTTGTGTGATC	15
10	TRBV16	CCTTTTCTGTCTGGTTCTTTGCTTC	16
11	TRBV17	CTAGACTTCTTTGCTGTGTGATCTTC	17
12	TRBV19	ATGAACAAGTGGGTTTTCTGCTG	18
13	TRBV20	ATGTTMCTGCTTCTATTACTTCTGG	19
14	TRBV23	CACGGCTCATTTGCTATGTAG	20
15	TRBV24	CAAGACTGCTCTGCTGTGTAG	21
16	TRBV26	ATGGCTACAAGGCTCCTCTG	22
17	TRBV29	TTAGGCTCATCTCTGCTGTGG	23
18	TRBV30	GACATTCCTGCTACTTCTTTGG	24
19	TRBV31	TGCTGTACTCTCTCCTTGCC	25

Table 4.2. Primers used in the amplification of TCRβ genes

In the potential case that high throughput sequencing reveals previously unobserved variable region alleles, the designed primer set may require additional members. However, the low Tm of the set is expected to allow for some mispairing. In addition, a relatively high degree of conservation is seen among known leader sequences. In cases where we have targeted a single gene with a degenerate primer from the set, no bias toward a particular conformation of that primer was detected (data not shown).

Primers were validated by amplification of murine cDNA. Lymph nodes were isolated from unimmunized mice, and individual reactions for each primer

were used to amplify α (Figure 4.2) and β (Figure 4.3) TCR genes. Amplification of genes was observed for each primer, with the exception of TRBV23. This gene has been previously noted as non-functional due to the absence of a splice donor site required for removal of the leader exon (183). Although proper-length transcripts have been identified for this gene, these may be the result of chromosomal rearrangement with a sister chromosome (184). Random individual bands from α and β were cloned into a topoisomerase cloning vector and sequenced to confirm the presence of at least one targeted gene amplified by each primer.

Figure 4.2. TRAV Primer Validation.

TCRα variable region fragments were amplified from cDNA of lymph node origin from unimmunized mice by individual primers from the primer set.

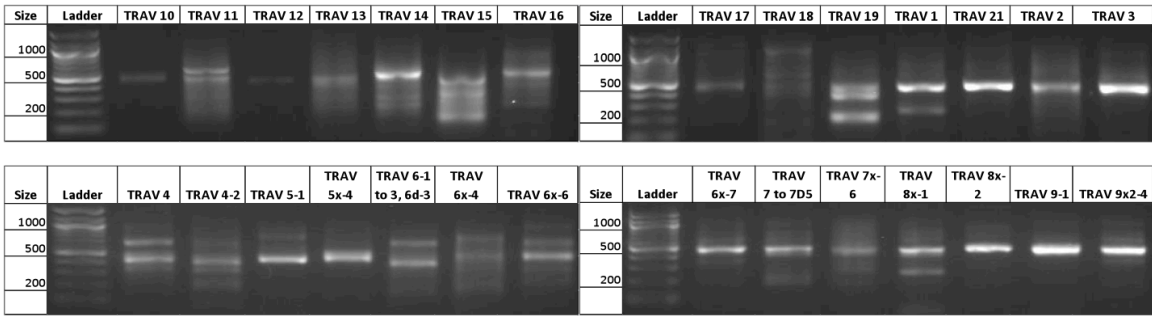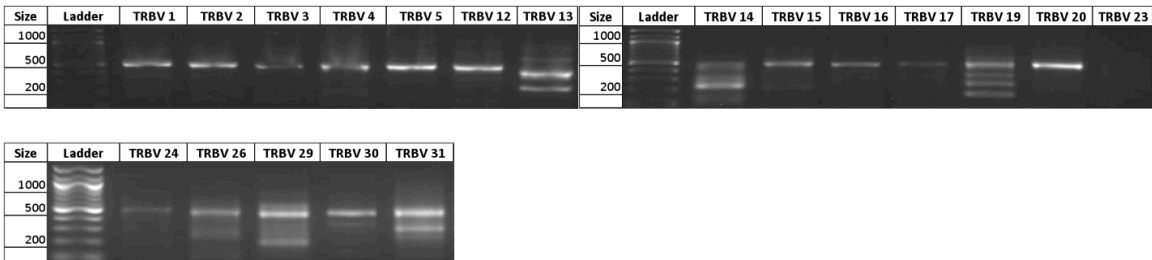
Figure 4.3. TRBV Primer Validation.

TCRβ variable region fragments were amplified from cDNA of lymph node origin from unimmunized mice by individual primers from the primer set.

4.3.2. Type B T Cell Receptor Isolation

As an initial feasibility study, we aimed to identify the TCR genes expreseed by several "Type B" T cell lines. These two lines are reactive to human myelin basic protein fragment 82-100 presented by HLA-DRB5*0101 were previously isolated from transgenic mice expressing this human MHC gene (185). The number of unique T cells within each line was previously unknown. First, total RNA was isolated from approximately ten million cells and amplified using the TCR primer set. Gel-purified PCR product was Topo-cloned and sequenced, to reveal a single functional α and β gene for the 4C4 line, using segments TRAV7D-2, TRAJ5; TRBV4, TRBD1 and TRBJ2-7.

For the 3B6 line, a single α gene using the TRAV7D-2 and TRAJ2 segments was recovered along with two functional β genes using the TRBV17, TRBD1, TRBJ1-1; TRBV20, TRBD1 and TRBJ1-4 segments and an additional non-functional gene using TRBV5, TRBD1 and TRBJ1-2. The results of the PCR amplification using individual β primers and 3B6 mRNA is shown in Figure 4.4. A splicing variant of the TRBV20 gene including the Cβ0 alternate cassette exon was also found, with an additional 24 amino acids inserted between the constant and variable domains. This variant has been previously demonstrated as functional as protein and occurring frequently in TRBJ1-containing receptors (186). The presence of three β genes implies the existence of at least one additional alpha gene for 3B6, but no others were recovered from traditional Sanger sequencing of Topo-cloned PCR products.
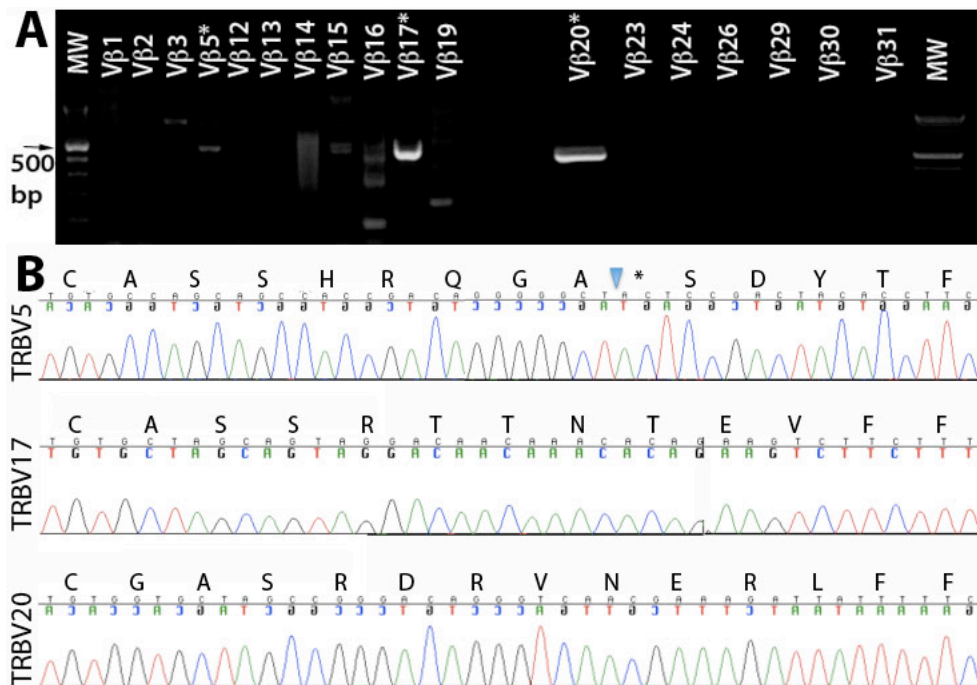
Figure 4.4. Amplification of 3B6 β genes.

cDNA from a "Type B" T cell line was amplified from using individual primers from the TCRβ set. One non-function and two functional rearranged genes were identified.

4.3.3. High Throughput Sequencing

A large-scale, high throughput sequencing project was designed using the primer set to analyze usage patterns in ovalbumin immunized versus control mice. CD4$^+$ and CD8$^+$ were separated from mouse splenocytes fourteen days after injection for two ova-immunized mice as well as two controls. Sequences were separated by mouse identity as well as T cell type. As of this writing, over 460,000 raw sequences had been generated at an average of around 320 bp in length with a mode of 420 bp. Analysis of this data is currently in progress.

Figure 4.5. Mouse CD8 TCRs for 454 Sequencing.

TCRs were amplified from the cDNA of CD8 isolated from splenocytes of an OVA-immunized mouse, separated on a 1.5% agarose gel (shown), purified, and bar-coded for 454 high-throughput sequencing. *Left*:TCRα, *Middle*: Ladder, *Right*,: TCRβ. The lower band in β is a result of the framework-targeting TRBV13 primer

A small-scale, preliminary trial was performed prior to this larger run. Here, 4,882 sequences were obtained from splenic CD4$^+$ TCR 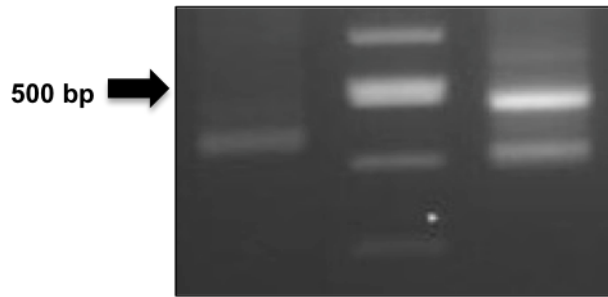cDNA amplified from a single, ovalbumin immunized mouse. Samples were first processed through IMGT's Vquest to remove non-TCR sequences, then filtered using crude cut-offs for length and sequence quality. Within the remaining 3,922 transcripts, Fifty-five of the 82 targeted genes were present, with notably high levels of diverse, functional transcripts from the TRAV7, 12, and 13 families (data not shown).

Germline segments not represented in the sequencing data were: TRAV1, TRAV2, TRAV5-1, TRAV5X-4, TRAV6x-6, TRAV8X-1, TRAV8X-2, TRAV10X, TRAV16X, TRAV17, TRAV18, TRAV19, and TRAV21. The TRAV 5X-4 primer sequence has previously been demonstrated to be functional in TCR amplification from a hybridoma (data not shown). Two other of primers may be targeting non-functional alleles. Variable region 18 (TRAV18) was targeted despite its one recorded allele being noted as an open reading frame (Figure). Likewise, the first recorded allele of variable region 19 is an open reading frame, and the second contains a stop codon, whereas a third reported allele was functional. Pseudogenes have also been recorded for alleles in both of the genes (TRAV10, TRAV10D) targeted by the TRAV10 primer.

Usage of TCR germline segments in the α region has historically been difficult to judge because of the sheer number of genes involved, and the lack of specific antibodies. The 13 primers that did not amplify TCRs in our preliminary sequencing experiment target a total of 17 genes. This is much lower than the average for our 28 member primer set, which targets a total of 82 unique genes. It is possible that these are infrequently used genes that have, as a result, not been selected for in the mouse like the larger-member families. This trend will be explored when a larger data set utilizing more refined filters becomes available.

The preliminary sequencing also revealed a possible error in the IMGT database, as TRAJ47 showed a conserved phenylalanine at position 117, rather than the denoted cysteine, and was observed in multiple functional transcripts.

CHAPTER 5

FUTURE DIRECTIONS

The various topics combined in this work may be synthesized in future research to yield a clearer picture of epitope presentation in health, disease, or vaccination. Several applications seem feasible. In the case of high-throughput sequencing, OVA 323-339 reactive subsets should be isolated prior to sequencing with newly- developed pMHC tetramers (97) designed to restrict binding of the peptide to the register of interest. By itself, the pre-screening step would yield insight into the nature of the TCR-based response to the antigen. The development of TCR-based reagents from sequenced receptors, however, requires additional information about chain pairing. A potential solution could be attempted based on the protein expression results observed here. Given the stability of TRBV13-2, clones from the sequence results containing this chain, or DO11β itself, might be used as a scaffold for pairing with frequently occurring α - chains from the sequence results. TRBV13-2 plus "high transcript-level α"- scTCRs could be grown in *E. coli* 96-well plates for rapid screening of tetramer or KJ1-26 binding. Used as a general platform, this method could effectively filter the process of identifying TCRs of interest, while simultaneously increasing the chances that these TCRs will be amenable to engineering in the robust bacterial host.

From a broader perspective, the power of display technologies has yet to be tapped for the engineering of T cell receptors. TRBV13-2 displays a degree of generality in expression when paired with various alpha chains, but remains the exception in this regard. Importantly, solubility-increasing amino acid substitutions identified for the 2C TCR (187) have the same effect when introduced onto many other Vβ chains, demonstrating that these protein can exhibit some generality. More work needs be done in exploring the structural

requirements for soluble expression. So that these can substitutions can be predicted a prior for additional TCR genes, without requiring directed evolution or crystallization experiments. DNA shuffling of related TCRs with wide ranges of bacterial expression levels could yield insight into important residues in solubility and stability in these molecules. Simple selection with a c-terminal epitope tag in phage display could be used in screening for expression.

Regarding the affinity-matured 176 variant, we are interested in the limits of KJ1-26 as a pMHC mimic. At some point, affinity with the antibody would likely fail to translate to the cognate antigen. A first step would be the investigation of the variant 174, which showed increased KJ1-26 binding as a soluble protein. Recombination of substitutions observed in 235, 176 and 174 would likely produce an "improved" variant, given the small coverage of six possible randomized codons that a 10^6-member library provides. CDR2β randomization of DO11 would also provide an interesting test case, given that others have shown that randomization of this MHC-targeted region can provide peptide specificity (149).

The 176 variant itself is also in need of further investigation. The peptide showed peptide specificity to OVA 323-339 in flow cytometry, but this could conceivably be due to the conformation of MHC induced by the presence of that peptide, which is commonly observed as differential mobility in an SDS-PAGE gel. Early experiments with maturation make it clear that cross-reactivity is a concern when targeting peptide specificity (188). Testing for cross-reactivity the QAV13-25 hen egg lysozyme variant peptide, in particular, would be of interest, given the evidence presented in Chapter 2 suggesting that this peptide binds to the MHC in similar fashion to OVA at the *N*-terminus.

Last, another potentially significant finding reported here is that flanking residues can increase peptide presentation. If this holds for other peptides, it may allow for engineering of peptides in vaccines in positions that don't affect T cell

interaction. This could be impactful, given that typical anchor residue substitution techniques can significantly impact T cell response.

BIBLIOGRAPHY

1. Alberts B*, et al.* (2002) *Molecular Biology of the Cell* (Garland Science, New York, NY) 4th Ed.
2. Alberts B (2002) *Molecular biology of the cell* (Garland Science, New York) 4th Ed pp xxxiv, 1548 p.
3. Janeway CA (2005) *Immunobiology: The Immune System in Health and Disease* (Garland Science Publishing, New York, NY) 6th Ed p 823.
4. McFarland BJ & Beeson C (2002) Binding interactions between peptides and proteins of the class II major histocompatibility complex. *Medicinal Research Reviews* 22(2):168-203.
5. http://nature.chem.msu.su/ (2005) V(D)J Recombination.
6. http://www.ahabs.wisc.edu/ (2004) Thymus.
7. Kyewski B & Derbinski J (2004) Self-representation in the thymus: An extended view. *Nature Reviews Immunology* 4(9):688-698.
8. Adorini L (2004) Immunotherapeutic approaches in multiple sclerosis. *Journal of the Neurological Sciences* 223(1):13-24.
9. Sasaki S, Takeshita F, Okuda K, & Ishii N (2001) Mycobacterium leprae and leprosy: A compendium. *Microbiology and Immunology* 45(11):729-736.
10. Faria AM & Weiner HL (2005) Oral tolerance. *Immunol Rev* 206:232-259.
11. Schwartz RH (2003) T cell anergy. *Annual Review of Immunology* 21:305-334.
12. Bluestone JA & Tang QZ (2004) Therapeutic vaccination using CD4(+) CD25(+) antigen-specific regulatory T cells. *Proceedings of the National Academy of Sciences of the United States of America* 101:14622-14626.
13. Asano M, Toda M, Sakaguchi N, & Sakaguchi S (1996) Autoimmune disease as a consequence of developmental abnormality of a T cell subpopulation. *Journal of Experimental Medicine* 184(2):387-396.
14. Randolph DA & Fathman CG (2006) Cd4+Cd25+ regulatory T cells and their therapeutic potential. *Annu Rev Med* 57:381-402.
15. Bluestone JA (2005) Regulatory T-cell therapy: is it ready for the clinic? *Nat Rev Immunol* 5(4):343-349.
16. Shevach EM (2009) Mechanisms of foxp3+ T regulatory cell-mediated suppression. *Immunity* 30(5):636-645.
17. Relland LM*, et al.* (2009) Affinity-based selection of regulatory T cells occurs independent of agonist-mediated induction of Foxp3 expression. *J Immunol* 182(3):1341-1350.
18. Jordan MS*, et al.* (2001) Thymic selection of CD4+CD25+ regulatory T cells induced by an agonist self-peptide. *Nature immunology* 2(4):301-306.
19. Hsieh CS*, et al.* (2004) Recognition of the peripheral self by naturally arising CD25+ CD4+ T cell receptors. *Immunity* 21(2):267-277.

20. Bautista JL*, et al.* (2009) Intraclonal competition limits the fate determination of regulatory T cells in the thymus. *Nature immunology* 10(6):610-617.

21. Rudolph MG & Wilson IA (2002) The specificity of TCR/pMHC interaction. *Curr Opin Immunol* 14(1):52-65.

22. Garcia KC, Adams JJ, Feng D, & Ely LK (2009) The molecular basis of TCR germline bias for MHC is surprisingly simple. *Nature immunology* 10(2):143-147.

23. Wu LC, Tuot DS, Lyons DS, Garcia KC, & Davis MM (2002) Two-step binding mechanism for T-cell receptor recognition of peptide MHC. *Nature* 418(6897):552-556.

24. Valitutti S, Muller S, Cella M, Padovan E, & Lanzavecchia A (1995) Serial triggering of many T-cell receptors by a few peptide-MHC complexes. *Nature* 375(6527):148-151.

25. Rabinowitz JD, Beeson C, Lyons DS, Davis MM, & McConnell HM (1996) Kinetic discrimination in T-cell activation. *Proceedings of the National Academy of Sciences of the United States of America* 93(4):1401-1405.

26. van der Merwe PA & Dushek O (2011) Mechanisms for T cell receptor triggering. *Nat Rev Immunol* 11(1):47-55.

27. Reichert JM (2008) Monoclonal antibodies as innovative therapeutics. *Curr Pharm Biotechnol* 9(6):423-430.

28. Kohler G & Milstein C (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *J Immunol* 174(5):2453-2455.

29. Morrison SL, Johnson MJ, Herzenberg LA, & Oi VT (1984) Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proceedings of the National Academy of Sciences of the United States of America* 81(21):6851-6855.

30. Jones PT, Dear PH, Foote J, Neuberger MS, & Winter G (1986) Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature* 321(6069):522-525.

31. Marks JD*, et al.* (1991) By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J Mol Biol* 222(3):581-597.

32. Knappik A*, et al.* (2000) Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *Journal of Molecular Biology* 296(1):57-86.

33. Cohen CJ*, et al.* (2003) Direct phenotypic analysis of human MHC class I antigen presentation: Visualization, quanitititation, and in situ detection of human viral epitopes using peptide-specific, MHC-restricted human recombinant antibodies. *Journal of Immunology* 170(8):4349-4361.

34. Andersen PS*, et al.* (1996) A recombinant antibody with the antigen-specific, major histocompatibility complex-restricted specificity of T cells. *Proceedings of the National Academy of Sciences of the United States of America* 93(5):1820-1824.

35. Porgador A, Yewdell JW, Deng YP, Bennink JR, & Germain RN (1997) Localization, quantitation, and in situ detection of specific peptide MHC class I complexes using a monoclonal antibody. *Immunity* 6(6):715-726.

36. Puri J, Arnon R, Gurevich E, & Teitelbaum D (1997) Modulation of the immune response in multiple sclerosis - Production of monoclonal antibodies specific to HLA/myelin basic protein. *Journal of Immunology* 158(5):2471-2476.

37. Engberg J, Krogsgaard M, & Fugger L (1999) Recombinant antibodies with the antigen-specific, MHC restricted specificity of T cells: novel reagents for basic and clinical investigations and immunotherapy. *Immunotechnology* 4(3-4):273-278.

38. Boulter JM & Jakobsen BK (2005) Stable, soluble, high-affinity, engineered T cell receptors: novel antibody-like proteins for specific targeting of peptide antigens. *Clinical and Experimental Immunology* 142(3):454-460.

39. de Haard HJ*, et al.* (1999) A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *J Biol Chem* 274(26):18218-18230.

40. Cohen CJ*, et al.* (2003) Direct phenotypic analysis of human MHC class I antigen presentation: visualization, quantitation, and in situ detection of human viral epitopes using peptide-specific, MHC-restricted human recombinant antibodies. *J Immunol* 170(8):4349-4361.

41. Andersen PS*, et al.* (1996) A recombinant antibody with the antigen-specific, major histocompatibility complex-restricted specificity of T cells. *Proceedings of the National Academy of Sciences of the United States of America* 93(5):1820-1824.

42. Chames P, Hufton SE, Coulie PG, Uchanska-Ziegler B, & Hoogenboom HR (2000) Direct selection of a human antibody fragment directed against the tumor T-cell epitope HLA-A1-MAGE-A1 from a nonimmunized phage-Fab library. *Proceedings of the National Academy of Sciences of the United States of America* 97(14):7969-7974.

43. Hulsmeyer M*, et al.* (2005) A major histocompatibility complex-peptide-restricted antibody and t cell receptor molecules recognize their target by distinct binding modes: crystal structure of human leukocyte antigen (HLA)-A1-MAGE-A1 in complex with FAB-HYB3. *J Biol Chem* 280(4):2972-2980.

44. Mareeva T, Martinez-Hackert E, & Sykulev Y (2008) How a T cell receptor-like antibody recognizes major histocompatibility complex-bound peptide. *J Biol Chem* 283(43):29053-29059.

45. Weiss GA, Watanabe CK, Zhong A, Goddard A, & Sidhu SS (2000) Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proceedings of the National Academy of Sciences* 97:8950-8954.

46. Murase K*, et al.* (2003) EF-Tu binding peptides identified, dissected, and affinity optimized by phage display. *Chemistry and Biology* 10:161-168.
47. Fazelinia H, Cirino PC, & Maranas CD (2007) Extending iterative protein redesign and optimization (IPRO) in protein library design for ligand specificity. *Biophysical Journal* 92:2120-2130.
48. Petrounia IP & Arnold FH (2000) Designed evolution of enzymatic properties. *Current Opinion in Biotechnology* 11:325-330.
49. Rubin-Pitel SB & Zhao H (2006) Recent advances in biocatalysis by directed enzyme evolution. *Combinatorial Chemistry and High Throughput Screening* 9:247-257.
50. Yu H, Tyo K, Alper H, Klein-Marcuschamer D, & Stephanopoulos G (2008) A high-throughput screen for hyaluronic acid accumulation in recombinant *Escherichia coli* transformed by libraries of engineered sigma factors. *Biotechnology and Bioengineering* 101:788-796.
51. Alper H, Fischer C, Nevoigt E, & Stephanopoulos G (2005) Tuning genetic control through promoter engineering. *Proceedings of the National Academy of Sciences* 102:12678-12683.
52. Tao H & Cornish VW (2002) Milestones in directed enzyme evolution. *Current Opinion in Chemical Biology* 6:858-864.
53. Ang EL, Obbard JP, & Zhao H (2009) Directed evolution of aniline dioxygenase for enhanced bioremediation of aromatic amines. *Appl Microbiol Biotechnol* 81(6):1063-1070.
54. Sidhu SS & Weiss GA (2004) Chapter 2. Oligonucleotide-Directed Construction of Phage Display Libraries. *Phage Display: A Practical Approach*, eds Lowman HL & Clackson T (Oxford University Press), pp 27-41.
55. Harvey BR*, et al.* (2004) Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries. *Proceedings of the National Academy of Sciences of the United States of America* 101(25):9193-9198.
56. Boder ET & Wittrup KD (1997) Yeast surface display for screening combinatorial polypeptide libraries. *Nature biotechnology* 15(6):553-557.
57. Pepper LR, Cho YK, Boder ET, & Shusta EV (2008) A decade of yeast surface display technology: where are we now? *Comb Chem High Throughput Screen* 11(2):127-134.
58. Hanes J & Pluckthun A (1997) In vitro selection and evolution of functional proteins by using ribosome display. *Proceedings of the National Academy of Sciences of the United States of America* 94(10):4937-4942.
59. Zahnd C, Amstutz P, & Pluckthun A (2007) Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target. *Nat Methods* 4(3):269-279.
60. Fletcher G, Mason S, Terrett J, & Soloviev M (2003) Self-assembly of proteins and their nucleic acids. *J Nanobiotechnology* 1(1):1.

61.   Suttle CA (2005) Viruses in the sea. *Nature* 437(7057):356-361.
62.   Smith GP (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 228(4705):1315-1317.
63.   Smothers JF, Henikoff S, & Carter P (2002) Tech.Sight. Phage display. Affinity selection from biological libraries. *Science* 298(5593):621-622.
64.   Barbas CFI (2001) *Phage Display: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.).
65.   Krebber A*, et al.* (1997) Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *Journal of Immunological Methods* 201(1):35-55.
66.   Clackson T, Hoogenboom HR, Griffiths AD, & Winter G (1991) Making antibody fragments using phage display libraries. *Nature* 352(6336):624-628.
67.   Fromant M, Blanquet S, & Plateau P (1995) Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal Biochem* 224(1):347-353.
68.   Stemmer WPC (1994) DNA Shuffling by Random Fragmentation and Reassembly - in-Vitro Recombination for Molecular Evolution. *Proceedings of the National Academy of Sciences of the United States of America* 91(22):10747-10751.
69.   Zhao HM, Giver L, Shao ZX, Affholter JA, & Arnold FH (1998) Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nature Biotechnology* 16(3):258-261.
70.   Kunkel TA (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci U S A* 82(2):488-492.
71.   Braunagel M & Little M (1997) Construction of a semisynthetic antibody library using trinucleotide oligos. *Nucleic Acids Research* 25(22):4690-4691.
72.   Sidhu SS (2005) *Phage display in biotechnology and drug discovery* (CRC Press/Taylor & Francis, Boca Raton) pp xviii, 748 p.
73.   Sanger F, Nicklen S, & Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74(12):5463-5467.
74.   Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, & Nyren P (1996) Real-time DNA sequencing using detection of pyrophosphate release. *Analytical biochemistry* 242(1):84-89.
75.   Ronaghi M, Uhlen M, & Nyren P (1998) A sequencing method based on real-time pyrophosphate. *Science* 281(5375):363, 365.
76.   Margulies M*, et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437(7057):376-380.
77.   Rudolph MG, Stanfield RL, & Wilson IA (2006) How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* 24:419-466.

78. Castellino F, Zhong G, & Germain RN (1997) Antigen presentation by MHC class II molecules: invariant chain function, protein trafficking, and the molecular basis of diverse determinant capture. *Human immunology* 54(2):159-169.

79. Chang ST, Ghosh D, Kirschner DE, & Linderman JJ (2006) Peptide length-based prediction of peptide-MHC class II binding. *Bioinformatics* 22(22):2761-2767.

80. Bordner AJ & Mittelmann HD (2010) Prediction of the binding affinities of peptides to class II MHC using a regularized thermodynamic model. *BMC Bioinformatics* 11:41.

81. Jardetzky TS*, et al.* (1996) Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. *Proceedings of the National Academy of Sciences of the United States of America* 93(2):734-738.

82. Nielsen M, Lund O, Buus S, & Lundegaard C (2010) MHC class II epitope predictive algorithms. *Immunology* 130(3):319-328.

83. Godkin AJ, Davenport MP, Willis A, Jewell DP, & Hill AV (1998) Use of complete eluted peptide sequence data from HLA-DR and -DQ molecules to predict T cell epitopes, and the influence of the nonbinding terminal regions of ligands in epitope selection. *J Immunol* 161(2):850-858.

84. Godkin AJ*, et al.* (2001) Naturally processed HLA class II peptides reveal highly conserved immunogenic flanking region sequence preferences that reflect antigen processing rather than peptide-MHC interactions. *J Immunol* 166(11):6720-6727.

85. Nielsen M, Lundegaard C, & Lund O (2007) Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics* 8:238.

86. Nielsen M & Lund O (2009) NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics* 10:296.

87. Jones EY, Fugger L, Strominger JL, & Siebold C (2006) MHC class II proteins and disease: a structural perspective. *Nat Rev Immunol* 6(4):271-282.

88. Kaushansky N*, et al.* (2009) HLA-DQB1*0602 determines disease susceptibility in a new "humanized" multiple sclerosis model in HLA-DR15 (DRB1*1501;DQB1*0602) transgenic mice. *J Immunol* 183(5):3531-3541.

89. Bankovich AJ, Girvin AT, Moesta AK, & Garcia KC (2004) Peptide register shifting within the MHC groove: theory becomes reality. *Mol Immunol* 40(14-15):1033-1039.

90. Stadinski BD*, et al.* (2010) Diabetogenic T cells recognize insulin bound to IAg7 in an unexpected, weakly binding register. *Proceedings of the National Academy of Sciences of the United States of America* 107(24):10978-10983.

91. Fairchild PJ, Pope H, & Wraith DC (1996) The nature of cryptic epitopes within the self-antigen myelin basic protein. *Int Immunol* 8(7):1035-1043.

92. Seamons A*, et al.* (2003) Competition between two MHC binding registers in a single peptide processed from myelin basic protein influences tolerance and susceptibility to autoimmunity. *The Journal of experimental medicine* 197(10):1391-1397.

93. Maverakis E, Beech JT, Schneider S, & Sercarz EE (2008) Presentation of a determinant by MHC class II can be prevented through competitive capture by a flanking determinant on a multideterminant peptide. *J Autoimmun* 31(1):59-65.

94. Shimonkevitz R, Kappler J, Marrack P, & Grey H (1983) Antigen recognition by H-2-restricted T cells. I. Cell-free antigen processing. *The Journal of experimental medicine* 158(2):303-316.

95. McFarland BJ, Sant AJ, Lybrand TP, & Beeson C (1999) Ovalbumin(323-339) peptide binds to the major histocompatibility complex class II I-A(d) protein using two functionally distinct registers. *Biochemistry* 38(50):16663-16670.

96. Robertson JM, Jensen PE, & Evavold BD (2000) DO11.10 and OT-II T cells recognize a C-terminal ovalbumin 323-339 epitope. *J Immunol* 164(9):4706-4712.

97. Landais E*, et al.* (2009) New design of MHC class II tetramers to accommodate fundamental principles of antigen presentation. *J Immunol* 183(12):7949-7957.

98. Nelson CA, Roof RW, McCourt DW, & Unanue ER (1992) Identification of the naturally processed form of hen egg white lysozyme bound to the murine major histocompatibility complex class II molecule I-Ak. *Proceedings of the National Academy of Sciences of the United States of America* 89(16):7380-7383.

99. Lazarski CA*, et al.* (2005) The kinetic stability of MHC class II:peptide complexes is a key parameter that dictates immunodominance. *Immunity* 23(1):29-40.

100. Scott CA, Peterson PA, Teyton L, & Wilson IA (1998) Crystal structures of two I-Ad-peptide complexes reveal that high affinity can be achieved without large anchor residues. *Immunity* 8(3):319-329.

101. Martineau P, Guillet JG, Leclerc C, & Hofnung M (1992) Expression of heterologous peptides at two permissive sites of the MalE protein: antigenicity and immunogenicity of foreign B-cell and T-cell epitopes. *Gene* 113(1):35-46.

102. Krebber A*, et al.* (1997) Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *Journal of immunological methods* 201(1):35-55.

103. Dang LH, Lien LL, Benacerraf B, & Rock KL (1993) A mutant antigen-presenting cell defective in antigen presentation expresses class II MHC molecules with an altered conformation. *J Immunol* 150(10):4206-4217.

104. Russell HI, York IA, Rock KL, & Monaco JJ (1999) Class II antigen processing defects in two H2d mouse cell lines are caused by point mutations in the H2-DMa gene. *European journal of immunology* 29(3):905-911.

105. Lazarski CA, Chaves FA, & Sant AJ (2006) The impact of DM on MHC class II-restricted antigen presentation can be altered by manipulation of MHC-peptide kinetic stability. *The Journal of experimental medicine* 203(5):1319-1328.

106. Comeau SR, Gatchell DW, Vajda S, & Camacho CJ (2004) ClusPro: a fully automated algorithm for protein-protein docking. *Nucleic acids research* 32(Web Server issue):W96-99.

107. Chaves FA, Richards KA, Torelli A, Wedekind J, & Sant AJ (2006) Peptide-binding motifs for the I-Ad MHC class II molecule: alternate pH-dependent binding behavior. *Biochemistry* 45(20):6426-6433.

108. McFarland BJ & Beeson C (2002) Binding interactions between peptides and proteins of the class II major histocompatibility complex. *Med Res Rev* 22(2):168-203.

109. Vogt AB, Kropshofer H, & Hammerling GJ (1997) How HLA-DM affects the peptide repertoire bound to HLA-DR molecules. *Human immunology* 54(2):170-179.

110. Nanda NK & Sant AJ (2000) DM determines the cryptic and immunodominant fate of T cell epitopes. *The Journal of experimental medicine* 192(6):781-788.

111. Busch R*, et al.* (2005) Achieving stability through editing and chaperoning: regulation of MHC class II peptide binding and expression. *Immunological reviews* 207:242-260.

112. Liang MN, Beeson C, Mason K, & McConnell HM (1995) Kinetics of the reactions between the invariant chain (85-99) peptide and proteins of the murine class II MHC. *Int Immunol* 7(9):1397-1404.

113. Weaver JM & Sant AJ (2009) Understanding the focused CD4 T cell response to antigen and pathogenic organisms. *Immunol Res*.

114. Abiru N*, et al.* (2000) Dual overlapping peptides recognized by insulin peptide B:9-23 T cell receptor AV13S3 T cell clones of the NOD mouse. *J Autoimmun* 14(3):231-237.

115. Levisetti MG, Suri A, Petzold SJ, & Unanue ER (2007) The insulin-specific T cells of nonobese diabetic mice recognize a weak MHC-binding segment in more than one form. *J Immunol* 178(10):6051-6057.

116. Kersh GJ*, et al.* (2001) Structural and functional consequences of altering a peptide MHC anchor residue. *J Immunol* 166(5):3345-3354.

117. Fremont DH, Hendrickson WA, Marrack P, & Kappler J (1996) Structures of an MHC class II molecule with covalently bound single peptides. *Science* 272(5264):1001-1004.

118. Wang JH*, et al.* (2001) Crystal structure of the human CD4 N-terminal two-domain fragment complexed to a class II MHC molecule. *Proceedings of the National Academy of Sciences of the United States of America* 98(19):10799-10804.

119. Sundberg EJ, Andersen PS, Schlievert PM, Karjalainen K, & Mariuzza RA (2003) Structural, energetic, and functional analysis of a protein-protein interface at distinct stages of affinity maturation. *Structure* 11(9):1151-1161.

120. Krogsgaard M*, et al.* (2003) Evidence that structural rearrangements and/or flexibility during TCR binding can contribute to T cell activation. *Mol Cell* 12(6):1367-1378.

121. McBeth C*, et al.* (2008) A new twist in TCR diversity revealed by a forbidden alphabeta TCR. *Journal of molecular biology* 375(5):1306-1319.

122. Gunther S*, et al.* (2010) Bidirectional binding of invariant chain peptides to an MHC class II molecule. *Proceedings of the National Academy of Sciences of the United States of America* 107(51):22219-22224.

123. Gervois N, Guilloux Y, Diez E, & Jotereau F (1996) Suboptimal activation of melanoma infiltrating lymphocytes (TIL) due to low avidity of TCR/MHC-tumor peptide interactions. *The Journal of experimental medicine* 183(5):2403-2407.

124. Parkhurst MR*, et al.* (1996) Improved induction of melanoma-reactive CTL with peptides from the melanoma antigen gp100 modified at HLA-A*0201-binding residues. *J Immunol* 157(6):2539-2548.

125. Bakker AB*, et al.* (1997) Analogues of CTL epitopes with improved MHC class-I binding capacity elicit anti-melanoma CTL recognizing the wild-type epitope. *Int J Cancer* 70(3):302-309.

126. Dyall R*, et al.* (1998) Heteroclitic immunization induces tumor immunity. *The Journal of experimental medicine* 188(9):1553-1561.

127. Overwijk WW*, et al.* (1998) gp100/pmel 17 is a murine tumor rejection antigen: induction of "self"-reactive, tumoricidal T cells using high-affinity, altered peptide ligand. *The Journal of experimental medicine* 188(2):277-286.

128. Valmori D*, et al.* (1999) Optimal activation of tumor-reactive T cells by selected antigenic peptide analogues. *Int Immunol* 11(12):1971-1980.

129. Slansky JE*, et al.* (2000) Enhanced antigen-specific antitumor immunity with altered peptide ligands that stabilize the MHC-peptide-TCR complex. *Immunity* 13(4):529-538.

130. Latek RR, Petzold SJ, & Unanue ER (2000) Hindering auxiliary anchors are potent modulators of peptide binding and selection by I-Ak class II

molecules. *Proceedings of the National Academy of Sciences of the United States of America* 97(21):11460-11465.

131. Kersh GJ, Kersh EN, Fremont DH, & Allen PM (1998) High- and low-potency ligands with similar affinities for the TCR: the importance of kinetics in TCR signaling. *Immunity* 9(6):817-826.

132. Cole DK*, et al.* (2010) Modification of MHC anchor residues generates heteroclitic peptides that alter TCR binding and T cell recognition. *J Immunol* 185(4):2600-2610.

133. Jiang W & Boder ET (2010) High-throughput engineering and analysis of peptide binding to class II MHC. *Proceedings of the National Academy of Sciences of the United States of America* 107(30):13258-13263.

134. Novotny J*, et al.* (1991) A soluble, single-chain T-cell receptor fragment endowed with antigen-combining properties. *Proceedings of the National Academy of Sciences of the United States of America* 88(19):8646-8650.

135. Kappler J, White J, Kozono H, Clements J, & Marrack P (1994) Binding of a soluble alpha beta T-cell receptor to superantigen/major histocompatibility complex ligands. *Proceedings of the National Academy of Sciences of the United States of America* 91(18):8462-8466.

136. Boulter JM*, et al.* (2003) Stable, soluble T-cell receptor molecules for crystallization and therapeutics. *Protein Engineering* 16(9):707-711.

137. Willcox BE*, et al.* (1999) Production of soluble alpha beta T-cell receptor heterodimers suitable for biophysical analysis of ligand binding. *Protein Science* 8(11):2418-2423.

138. Margulies DH (1997) Interactions of TCRs with MHC-peptide complexes: a quantitative basis for mechanistic models. *Curr Opin Immunol* 9(3):390-395.

139. Weber KS, Donermeyer DL, Allen PM, & Kranz DM (2005) Class II-restricted T cell receptor engineered in vitro for higher affinity retains peptide specificity and function. *Proceedings of the National Academy of Sciences of the United States of America* 102(52):19033-19038.

140. Shusta EV, Holler PD, Kieke MC, Kranz DM, & Wittrup KD (2000) Directed evolution of a stable scaffold for T-cell receptor engineering. *Nature Biotechnology* 18(7):754-759.

141. Holler PD*, et al.* (2000) In vitro evolution of a T cell receptor with high affinity for peptide/MHC. *Proceedings of the National Academy of Sciences of the United States of America* 97(10):5387-5392.

142. Holler PD, Chlewicki LK, & Kranz DM (2003) TCRs with high affinity for foreign pMHC show self-reactivity. *Nature Immunology* 4(1):55-62.

143. Chlewicki LK, Holler PD, Monti BC, Clutter MR, & Kranz DM (2005) High-affinity, peptide-specific T cell receptors can be generated by mutations in CDR1, CDR2 or CDR3. *Journal of Molecular Biology* 346(1):223-239.

144. Chervin AS, Aggen DH, Raseman JM, & Kranz DM (2008) Engineering higher affinity T cell receptors using a T cell display system. *Journal of immunological methods* 339(2):175-184.

145. Onda T, *et al.* (1995) A phage display system for detection of T cell receptor-antigen interactions. *Molecular Immunology* 32(17-18):1387-1397.

146. Weidanz JA, Card KF, Edwards A, Perlstein E, & Wong HC (1998) Display of functional alpha beta single-chain T-cell receptor molecules on the surface of bacteriophage. *Journal of Immunological Methods* 221(1-2):59-76.

147. Li Y, *et al.* (2005) Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nature Biotechnology* 23(3):349-354.

148. Haidar JN, *et al.* (2009) Structure-based design of a T-cell receptor leads to nearly 100-fold improvement in binding affinity for pepMHC. *Proteins* 74(4):948-960.

149. Dunn SM, *et al.* (2006) Directed evolution of human T cell receptor CDR2 residues by phage display dramatically enhances affinity for cognate peptide-MHC without increasing apparent cross-reactivity. *Protein science : a publication of the Protein Society* 15(4):710-721.

150. Xu XN & Screaton GR (2002) MHC/peptide tetramer-based studies of T cell function. *Journal of Immunological Methods* 268(1):21-28.

151. Laugel B, *et al.* (2005) Design of soluble recombinant T cell receptors for antigen targeting and T cell inhibition. *Journal of Biological Chemistry* 280(3):1882-1892.

152. Subbramanian RA, *et al.* (2004) Engineered T-cell receptor tetramers bind MHC-peptide complexes with high affinity. *Nature Biotechnology* 22(11):1429-1434.

153. Haskins K, *et al.* (1983) The major histocompatibility complex-restricted antigen receptor on T cells. I. Isolation with a monoclonal antibody. *The Journal of experimental medicine* 157(4):1149-1169.

154. Marrack P, Shimonkevitz R, Hannum C, Haskins K, & Kappler J (1983) The major histocompatibility complex-restricted antigen receptor on T cells. IV. An antiidiotypic antibody predicts both antigen and I-specificity. *The Journal of experimental medicine* 158(5):1635-1646.

155. Richman SA & Kranz DM (2007) Display, engineering, and applications of antigen-specific T cell receptors. *Biomolecular engineering* 24(4):361-373.

156. Boulter JM & Jakobsen BK (2005) Stable, soluble, high-affinity, engineered T cell receptors: novel antibody-like proteins for specific targeting of peptide antigens. *Clinical and experimental immunology* 142(3):454-460.

157. Stauss HJ*, et al.* (2007) Monoclonal T-cell receptors: new reagents for cancer therapy. *Molecular therapy : the journal of the American Society of Gene Therapy* 15(10):1744-1750.

158. Gesser B*, et al.* (1997) Identification of functional domains on human interleukin 10. *Proceedings of the National Academy of Sciences of the United States of America* 94(26):14620-14625.

159. Kurte M*, et al.* (2004) A synthetic peptide homologous to functional domain of human IL-10 down-regulates expression of MHC class I and Transporter associated with Antigen Processing 1/2 in human melanoma cells. *J Immunol* 173(3):1731-1737.

160. Lopez MN*, et al.* (2011) A synthetic peptide homologous to IL-10 functional domain induces monocyte differentiation to TGF-beta+ tolerogenic dendritic cells. *Immunobiology* 216(10):1117-1126.

161. Hayhurst A (2000) Improved expression characteristics of single-chain Fv fragments when fused downstream of the Escherichia coli maltose-binding protein or upstream of a single immunoglobulin-constant domain. *Protein expression and purification* 18(1):1-10.

162. Kieke MC*, et al.* (1999) Selection of functional T cell receptor mutants from a yeast surface-display library. *Proceedings of the National Academy of Sciences of the United States of America* 96(10):5651-5656.

163. Garcia KC, Radu CG, Ho J, Ober RJ, & Ward ES (2001) Kinetics and thermodynamics of T cell receptor- autoantigen interactions in murine experimental autoimmune encephalomyelitis. *Proceedings of the National Academy of Sciences of the United States of America* 98(12):6818-6823.

164. Jung S & Pluckthun A (1997) Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting. *Protein engineering* 10(8):959-966.

165. Al-Lazikani B, Lesk AM, & Chothia C (2000) Canonical structures for the hypervariable regions of T cell alphabeta receptors. *Journal of molecular biology* 295(4):979-995.

166. Maynard J*, et al.* (2005) Structure of an autoimmune T cell receptor complexed with class II peptide-MHC: insights into MHC bias and antigen specificity. *Immunity* 22(1):81-92.

167. Schatz PJ (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli. *Biotechnology (N Y)* 11(10):1138-1143.

168. Avidity I ((http://www.avidity.com/t-technologyfaq.aspx).

169. Rice JJ, Schohn A, Bessette PH, Boulware KT, & Daugherty PS (2006) Bacterial display using circularly permuted outer membrane protein OmpX yields high affinity peptide ligands. *Protein science : a publication of the Protein Society* 15(4):825-836.

170. Grussenmeyer T, Scheidtmann KH, Hutchinson MA, Eckhart W, & Walter G (1985) Complexes of polyoma virus medium T antigen and cellular

proteins. *Proceedings of the National Academy of Sciences of the United States of America* 82(23):7952-7954.

171. Abdullah A, McCauley RL, & Herndon DN (1991) Stimulation of human dermal fibroblasts with interleukin 2. *The Journal of burn care & rehabilitation* 12(1):23-25.

172. Kretzschmar T & Geiser M (1995) Evaluation of antibodies fused to minor coat protein III and major coat protein VIII of bacteriophage M13. *Gene* 155(1):61-65.

173. Malik P*, et al.* (1996) Role of capsid structure and membrane protein processing in determining the size and copy number of peptides displayed on the major coat protein of filamentous bacteriophage. *Journal of molecular biology* 260(1):9-21.

174. Sidhu SS, Weiss GA, & Wells JA (2000) High copy display of large proteins on phage for functional selections. *Journal of molecular biology* 296(2):487-495.

175. Weiss GA, Wells JA, & Sidhu SS (2000) Mutational analysis of the major coat protein of M13 identifies residues that control protein display. *Protein science : a publication of the Protein Society* 9(4):647-654.

176. Aggen DH*, et al.* (2011) Identification and engineering of human variable regions that allow expression of stable single-chain T cell receptors. *Protein engineering, design & selection : PEDS* 24(4):361-372.

177. Zehn D, Lee SY, & Bevan MJ (2009) Complete but curtailed T-cell response to very low-affinity antigen. *Nature* 458(7235):211-214.

178. Wang C*, et al.* (2010) High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proceedings of the National Academy of Sciences of the United States of America* 107(4):1518-1523.

179. Acha-Orbea H*, et al.* (1988) Limited heterogeneity of T cell receptors from lymphocytes mediating autoimmune encephalomyelitis allows specific immune intervention. *Cell* 54(2):263-273.

180. Lefranc MP*, et al.* (1999) IMGT, the international ImMunoGeneTics database. *Nucleic acids research* 27(1):209-212.

181. Brochet X, Lefranc MP, & Giudicelli V (2008) IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic acids research* 36(Web Server issue):W503-508.

182. Bosc N & Lefranc MP (2003) The mouse (Mus musculus) T cell receptor alpha (TRA) and delta (TRD) variable genes. *Developmental and comparative immunology* 27(6-7):465-497.

183. Currier JR, Deulofeut H, Barron KS, Kehn PJ, & Robinson MA (1996) Mitogens, superantigens, and nominal antigens elicit distinctive patterns of TCRB CDR3 diversity. *Human immunology* 48(1-2):39-51.

184. Bernardin F, *et al.* (2003) Estimate of the total number of CD8+ clonal expansions in healthy adults using a new DNA heteroduplex-tracking assay for CDR3 repertoire analysis. *Journal of immunological methods* 274(1-2):159-175.

185. Kawamura K, McLaughlin KA, Weissert R, & Forsthuber TG (2008) Myelin-reactive type B T cells and T cells specific for low-affinity MHC-binding myelin peptides escape tolerance in HLA-DR transgenic mice. *J Immunol* 181(5):3202-3211.

186. Aublin A, *et al.* (2006) A natural structural variant of the mouse TCR beta-chain displays intrinsic receptor function and antigen specificity. *J Immunol* 177(12):8587-8594.

187. Shusta EV, Holler PD, Kieke MC, Kranz DM, & Wittrup KD (2000) Directed evolution of a stable scaffold for T-cell receptor engineering. *Nature biotechnology* 18(7):754-759.

188. Holler PD, Chlewicki LK, & Kranz DM (2003) TCRs with high affinity for foreign pMHC show self-reactivity. *Nature immunology* 4(1):55-62.