REGULATION OF SELECTION AND CENTRAL TOLERANCE BY THE N-TERMINAL REGION OF RAG1

Thomas Niels Burn

A DISSERTATION

in

Immunology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2020

Supervisor of Dissertation

_____
Edward Malachi Behrens, M.D.
Joseph Hollander Associate Professor in Pediatric Rheumatology

Graduate Group Chairperson

_____
David Michael Allman, Ph.D.
Professor of Pathology and Laboratory Medicine

Dissertation Committee:
Martha S. Jordan (Chair), Ph.D., Research Associate Professor of Pathology and Laboratory Medicine

Laurence C. Eisenlohr, V.M.D., Ph.D., Professor of Pathology and Laboratory Medicine

Craig H. Bassing, Ph.D., Associate Professor of Pathology and Laboratory Medicine

Phillip Scott, Ph.D., Professor of Microbiology and Immunology

REGULATION OF SELECTION AND CENTRAL TOLERANCE BY THE N-TERMINAL REGION OF RAG1

DEDICATION

To Mum, Zane, Hamish, and Marshall, whom I have spent too long away from in pursuit of this degree. And to Clarice, who was there with me during some difficult times and has always been my biggest supporter.

ACKNOWLEDGEMENTS

It is of course cliché, but the list of people I need to acknowledge here is extensive, and definitely lacking, as this thesis and degree could not have been achieved without the help and support of an incredible number of people.

First, I have to acknowledge Ed, from whom I have learnt so much from and who is really an inspirational scientist and mentor. The path to this Ph.D. thesis was by no means regular, but Ed has been incredibly supportive in undertaking a project that is well outside either of our wheelhouses. I greatly admire his approach to science, and his willingness to tackle a variety of interesting scientific problems, while maintaining an incredibly taxing array of clinical duties. I feel as though we have very similar perspectives on life and science which has made working with him very easy. I have been very lucky to have been able to work on a number of different projects throughout my Ph.D. and it is somewhat unfortunate that this thesis cannot cover all of them. The Behrens Scientific Method is something I hope to forever utilize in my career going forward.

Everyone in the Behrens Lab has been extremely helpful to me over the years, but I cannot go past mentioning Julia Rood who, along with Nian Chu, taught me everything I needed to know in order to tackle this Ph.D., and has become a close friend and someone who I admire greatly.

Thank you to our many collaborators, joint lab meeting attendees, and my committee for the vast amount of help you all have given me in driving this thesis. Special mention must be made to Craig Bassing and his lab (especially Charline Miot) for this thesis truly has been a close collaboration between our labs.

The Penn IGG program has been an incredible group to be a part of, and this is largely due to the incredible efforts of Mary Taylor and Dave Allman; but also due to the people who I have ridden this sometimes-brutal rollercoaster with. The Class of 2014 or “Brohort” of Brenal, Tanner, Glendon, Omar, Mike, and Jon have made this a truly enjoyable experience. They are all incredibly talented yet extremely down-to-earth, and even though my liver is worse off for knowing them, this Ph.D. would have been miserable without them.

Thank you to my close friend Morris and his family (Jeanette, Pete, Anna, Will, Dan, and Lydia) who have adopted me and become my American family so that I’ve always had a place to go for Thanksgiving, Christmas, or just to escape. Likewise, to Clarice’s family who have always made me feel so welcome, and whose office I wrote most of this thesis in!

Thank you to Mum, and my brothers Zane, Hamish, and Marshall who I have been away from for so long. While we haven’t seen each other often, when we do it feels like last week. The same goes for my close friends Willie, Aaron, Phil, and Lachlan from back home in New Zealand. I am looking forward to being closer all of you in the coming months.

Last but not least, thank you to Clarice who has always had my back, and has been my biggest supporter in life (and on the rugby sideline!). I love you very much and cannot wait to move with you and start the next chapter of our life in Melbourne.

ABSTRACT

REGULATION OF SELECTION AND CENTRAL TOLERANCE BY THE N-TERMINAL REGION OF RAG1

Thomas Niels Burn

Edward Malachi Behrens

The RAG1/RAG2 complex rearranges antigen receptor gene segments during VDJ recombination to allow for the generation of a vast array of antigen receptors with specificities against a virtually infinite number of insults. Direct consequences of VDJ recombination include the formation of non-functional receptors, and receptors that are self-reactive. For $\alpha\beta$ T cells, mechanisms have evolved to delete cells bearing non-functional and self-reactive T cell receptors (TCRs) from the peripheral repertoire. These mechanisms rely on self-antigen:TCR interactions of intermediate affinity during development. However, the requirement for TCR selection on self-peptides results in a tenuous state where some autoreactive TCRs might not signal for deletion. A mechanism to limit this outcome is one where T cells have heightened TCR sensitivity during development such that the self-antigen they are selected on will no longer be stimulatory upon reencounter in the periphery. How this mechanism is regulated is unknown. The RAG1/RAG2 proteins have recently been shown to have functions in addition to VDJ recombination, including induction of a transcriptional program that is necessary for the expression of a number of lymphocyte-specific genes. We show that the N-terminal domain of RAG1 acts to control expression of a number of proteins during T cell development. Identification of a novel nonsense mutation in the N-terminal region of RAG1

lead us to identify a number of N-truncated isoforms that are made via internal translation initiation. Using a mouse with an inactivating mutation in the N-terminal E3 ligase domain of RAG1, we show that N-terminal RAG1 activity is essential for the upregulation of a number of TCR signaling molecules during development. In the absence of this activity, positive and negative selection are impaired, and mature T cells are hyper-active and cause disseminated pathology when transferred into $RAG1^{-/-}$ hosts. Correspondingly, mutations in the N-terminal, non-catalytic region of RAG1 are associated with autoimmune diseases. We propose a model where RAG1 acts as a biological clock to tune up TCR signaling during development; and hypothesize that N-truncated RAG1 isoforms can have regulatory roles in VDJ recombination and perhaps lead to heterogenous self-antigen reactivity within developing thymocytes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1: An adaptive immune response

The arms race between potentially deadly pathogens, and the organisms that they target, has driven the evolution of an extensive array of mechanisms by which species as genetically simple as bacteria, or as complicated as vertebrates, have developed to protect themselves. The 'immune systems' of these different classes of life must have three essential components in order to be effective mechanisms of defense:

1. They must be able to defend against a pathogen that they have never encountered before, nor have any pre-determined knowledge of its form.
2. They must be able to remember pathogens that they have previously encountered to prevent re-infection in the future.
3. They must specifically recognize only non-self, in order to prevent self-destruction.

Viruses, which are obligate parasites, are common enemies of both bacteria and vertebrates alike. Viruses that infect bacteria, bacteriophages, are one of the most common and diverse lifeforms on the planet (1-3), and bacteriophages remove approximately 20-40% of all bacterial cells from the oceans of the world each day (4). The mechanism by which bacteria develop immunity to such constant attacks is through a method that can be thought of as 'viral capture'. The CRISPR (clustered regularly interspaced short palindromic repeats) system that has gained extensive sex appeal in recent years due to its ability to precisely edit DNA in mammalian systems (5, 6), was evolved as a mechanism by many species of bacteria to recognize, capture, and store pieces of non-self, bacteriophage genomes, destroy the genome of the invading phage,

and rapidly defend against any future phage invasions with similar genetic information (7-10). Thus, CRISPR provides a beautifully elegant mechanism by which, with a relatively simple genome, bacteria can defend against viral infection in a specific and remembered fashion.

Instead of adopting a 'capture and destroy' approach to defense against unknown pathogens, jawed vertebrates have evolved an 'observe the protein universe but ignore self' approach. In the late 19th century, Emil von Behring and Shibasaburo Kitasato showed that transfer of serum from animals previously infected with tetanus and diphtheria was able to both prevent infection in uninfected recipients, as well as cure infected animals, thus displaying for the first time a system of both specific and remembered immune protection, garnering the first Nobel Prize in Medicine (11, 12). Not long after, Paul Ehrlich proposed a theory, known as the "Side-Chain Theory", which stated that specific receptors present on cell surfaces can, upon encounter with a specific 'toxin', be cast off to act as soluble 'anti-toxin' (13, 14). This theory, also worthy of the Nobel Prize, was the basis for the clonal selection theories of Jerne and Burnet, 60 years later (15, 16). Burnet's clonal selection theory is the presently accepted theory. Briefly, the clonal selection theory, and the basis for the adaptive immune system of jawed vertebrates, is such that cells circulate throughout the body expressing an exceptionally diverse array of receptors capable of observing the entire universe of prospective antigen. When an individual cell encounters an antigen it can 'see', extensive expansion of this 'clone' occurs, and a series of events unfold to see eliminate the offending antigen. Such cells, classed 'lymphocytes' express on their surfaces, highly specific antigen receptors and are divided into two varieties: B cells, which upon activation secrete the soluble effectors

described by von Behring and Shibasaburo, known as antibodies; and T cells which act as field generals to coordinate the immune response through production of cytokines, and can kill infected cells to stop the spread of intracellular pathogens.

The major problem of the clonal selection theory is that is seemingly impossible to encode in the genome of a human, enough receptors in order to protect against the universe of possible invading pathogens. The extent of this problem is evident within the following numbers:

- The human genome: ~$3x10^9$ base pairs (bp) (17)
- Length of cDNA to encode a T cell receptor: ~1200bp (TCRβ only)(18)
- Estimated number of naïve T cells in an adult human: $3x10^{11}$ cells (19)
- Estimated number of unique TCRβ sequences in an adult human: $1x10^6$ (20) to $1x10^8$ (21)
- DNA required to encode 10^6-10^8 different TCRβ's: ~$1x10^9$ to ~$1x10^{11}$ bp, i.e. 40% of, to 40x the total human genome.

Even the lower end of this estimate is unfeasible given only ~1% of the human genome is protein coding, (~$3x10^7$ bp), and encodes for ~21000 proteins, far less than the expected number of different TCRβ chains expected to be present in an adult human (10^6-10^8). Additionally, these figures only account for one chain of the T cell receptor (discussed in more detail further along), and do not include B cell receptors (with an estimated 10^9 unique receptors (22)). Luckily, evolution has garnered humans with a mechanism that can generate in excess of 10^{15} different T cell receptors (19, 23), that are encoded by less

than 2.5Mb of genomic DNA, less than 0.1% of the total human genome (combined length of TCRβ and TCRα loci - NCBI).

1.2: Development of T cell receptor diversity

The development of antigen receptor diversity in humans and other vertebrates involves the somatic rearrangement of germline encoded DNA segments that was first described for the immunoglobulin heavy chain (IgH, the heavy chain of the B cell receptor) (24-26). Subsequently, the T cell receptor was shown to rearrange in similar fashion (18, 27, 28). Formation of the B and T cell receptor via this mechanism is known as "VDJ recombination".

The structure of the T cell receptor is such that it contains two chains, either TCRα paired with TCRβ, or TCRγ paired with TCRδ. The majority of T cells within humans and mice, contain αβ TCRs and these T cells are the focus of this dissertation. The locus of TCRβ consists of the following elements (from 5' to 3') (29):

- "Variable" (V) gene segments (n=30)
- "Diversity" (D) gene segments (n=2)
- "Joining" (J) gene segments (n=12, 6 adjacent to each D segment)
- Constant (C) gene segment

In order to generate diversity, a D segment is rearranged with a J segment, before one of the V segments rearranges to this DJ combination. Each of these recombination events either deletes out, or inverts the intervening DNA (29). The TCRα locus (which serves also as the TCRδ locus) contains similar elements but lacks D segments hence only requires

a V→J rearrangement. Likewise, heavy (IgH) and light (Igκ or Igλ) chains recombine similarly for construction of the B cell receptor. The pairing of different V regions with different D and J regions provides "combinatorial diversity", but further diversity comes from the variable addition or subtraction of nucleotides between each of the recombined regions, termed "junctional diversity" (26, 30). Finally, different TCRβ chains can pair with different TCRα chains to form the fully rearranged T cell receptor. Hence, with a relatively small amount of genetic material, an incredible amount of diversity can be achieved between TCRs of different T cells.

1.3: The VDJ recombination reaction

Flanking the V, D, and J gene segments are conserved DNA sequences termed 'recombination signal sequences' or RSS's. These RSS's consist of heptamer and nonamer sequences that are separated by either a 12 or 23 bp spacer sequence (29, 31). The recombination activating gene complex (RAG – discussed in more detail in section 1.4 below), recognizes RSSs flanking gene segments that are to be recombined, and brings them in close proximity to each other (32, 33). The 12/23 rule, which requires that an RSS with a 12bp spacer is recombined with an RSS containing a 23bp spacer, provides control over which segments are permitted to recombine, and in which direction (34-36). For example, the 3' end of V_β segments contain 23bp RSSs and the 5' end of D_β segments contain 12bp RSSs, thus ensuring the correct 5' to 3' orientation of V→DJ (37).

The RAG complex contains endonuclease activity, such that after it has bound two RSSs, and brought them in close proximity (mechanisms of which will not be discussed here), a single stranded DNA cut is made between the RSS and the relevant gene segments (i.e. V, D, or J). At this point, the RAG complex, along with HMGB1 or HMGB2

(38), holds in close proximity the two coding ends (the ends of the V or D or J segments), and two signal ends (the ends with the RSS sites) that terminate in covalently sealed hairpins (39). The hairpins are opened by the proteins Artemis and DNA-PKcs (40), before the free DNA ends are ligated together via non-homologous end joining (NHEJ) that utilizes the proteins Ku70, Ku80, XRCC4, DNA Ligase IV, and XLF (39, 41). The precise position of the opening of the hairpin by the Artemis:DNA-PKcs complex can vary from 1-4 nucleotides from the tip of the hairpin (42), and at this stage, the enzyme TdT can add nucleotides to these junctions (43). Both of these mechanisms contribute to the junctional diversity between ligated gene segments that was discussed earlier. The orientation in which the RAG and HMGB proteins hold the DNA ends results in the formation of two DNA joins; a coding join that links a V to a D, or a D to a J segment, and a signal join linking two RSSs together. The signal join is either spliced out as an excision circle or inverted depending on the orientation of the sequences being joined.

1.4: The recombination activating gene (RAG) complex

Key in this reaction, and essential for the production of T cell and B cell receptors, and thus adaptive immunity, is the RAG complex. The RAG complex is a hetero-tetramer consisting of two RAG1, and two RAG2 proteins. Both RAG1 and RAG2 are essential for formation of functional antigen receptors, and absence of these genes leads to T and B cell severe combined immunodeficiency (SCID) (44, 45). Indeed, ‘domestication’ of the RAG complex from a transposon was key to the evolution of this strategy of antigen receptor rearrangement, and adaptive immunity that was adopted by jawed vertebrates (46).

RAG1 contains the catalytic domain that is necessary for DNA cleavage, but RAG1 and RAG2 are both required for any catalytic activity to occur. The focus of this thesis lies with RAG1 and hence the following discussion will be centered on the structure and function of RAG1. RAG1 is a protein of 1040 amino acids (mouse, human = 1043) and consists of 'core' and 'non-core'/'N-terminal domains' (Figure 1-1). The core domain contains the minimally essential components that are required for VDJ recombination, including RAG2 binding, nonamer and heptamer binding, and endonuclease domains (39), yet in $RAG1^{core}$ mice, VDJ recombination is impaired (47, 48) suggesting N-terminal regions of RAG1 have roles in VDJ recombination and T cell development. Indeed, the RAG1 N-terminus contains an E3 ubiquitin ligase domain that monoubiquitylates histones to enhance RAG cleavage (49-53); binds a kinase that helps repair RAG double stranded DNA breaks (DSBs) (54); and associates with another ubiquitin ligase, VprBP, and an RNA splicing factor for which potential functions of these interactions remain unknown (54-57). The goal of this discussion is to introduce the idea that RAG1 may have functions in addition to its key role in VDJ recombination.

1.5: αβ-T cell development and central tolerance

Now that I have introduced the mechanisms and some of the key proteins behind the generation of a diverse antigen receptor repertoire, a mechanism that can theoretically make $>10^{15}$ different receptors from ~2.5Mb of genomic DNA, I will acknowledge the sharp blade on the other side of this powerful sword. In making such a diverse repertoire of receptors, it is with high probability that many of these created receptors will be specific to 'self' and thus can cause autoimmune reactions. Additionally, because of the imprecise nature of annealing at the signal joins, two-thirds of VDJ rearrangements will be out of

frame, and lead to no functional TCR expression. The thymus is the specialized organ in which T cells develop (58). Humans with genetic abnormalities that lead to the absence of a thymus, such as in DiGeorge Syndrome/22q11.2 deletion syndrome, have severely impaired T cell development and T cell lymphopenia (59).

The thymus serves two major functions to address the issues stated above:

1. It ensures that T cells without a functional TCR are removed from the peripheral repertoire.
2. It ensures that T cells with self-reactive/autoreactive TCRs are removed from the peripheral repertoire.

T cells develop from hematopoietic stem cells (HSCs) from the bone marrow. Within the bone marrow HSCs can be directed to gain lymphocyte and lose myeloid differentiation capacity to become the common lymphoid progenitor (CLP) (60). CLPs contain B cell, T cell, and NK cell differentiation capacity (60). CLPs can then migrate to the thymus and enter the cortical region as early thymic progenitors (ETP) or DN1 (double negative 1 – as determined by absence of CD4 and CD8 surface expression) cells (Figure 1-2 A). These cells which are $CD44^{hi}cKIT^{+}CD25^{-}$ then upregulate CD25, and down-regulate CD44/cKIT to become DN3 cells (61)(Figure 1-2 A). At this stage, TCRβ is rearranged. In order for further developmental progression, successful rearrangement of the TCRβ chain, pairing of this TCRβ with pre-Tα to form the pre-TCR, and pre-TCR signaling is required. Upon pre-TCR signaling, RAG expression is suppressed which is essential to prevent recombination of the other TCRβ allele (62, 63), and hence the possibility of dual TCRβ expression. In the absence of the pre-TCR signal due to an out-of-frame rearrangement,

RAG can continue to rearrange the second allele of TCRβ to improve the odds of the formation of a functional TCRβ (63). No ligand is thought to be necessary for pre-TCR signaling (64, 65). Pre-TCR signaling also induces a proliferative burst and cells progress to the DP stage of development (66) (Figure 1-2 B).

At the DP developmental stage, the majority of the cells exist as $CD69^{-}CCR7^{-}$, Stage 1 cells (Figure 1-2 C), and here is when TCRα is rearranged. Upon successful rearrangement of a functional TCRα chain and pairing with TCRβ, cells get a TCR signal if they can successfully bind peptide:major histocompatibility complexes (MHC) on the surface of cortical thymic epithelial cells (cTECs) or cortical dendritic cells (cDCs), a process known as “positive selection” (67). Positive selection results in upregulation of CD69 (Figure 1-2 C). Three important things to note here:

- Approximately 50million DP thymocytes are produced each day in a mouse, and 90% of these will die by ‘neglect’ because they were unable to successfully rearrange TCRα, or they made a useless TCR (67, 68).
- A useful TCR must detect the combination of peptide and MHC. MHC molecules present short peptides that originate from within the presenting cell (69). For functional TCR signaling with an pMHC complex, a co-receptor must also bind (CD8 for MHCI, CD4 for MHCII). Since DP cells express both, positive selection can occur on either MHCI or II.
- By necessity, the peptide:MHC complexes that developing αβ T cells observe are presenting self-peptide, or at the very least altered self as cTECs have been shown to express a thymus-specific proteasome (69). The thymic proteasome degrades

proteins in the thymus in a slightly different manner than the normal proteasome, but even these proteasomes lead to expression of ~70% of commonly presented peptides (69).

Given the necessity for some degree of self-reactivity for successful positive selection, herein lies another problem: how does the thymus prevent escape of autoreactive cells? Successful positive selection of cells will subsequently upregulate CCR7 allowing migration of these cells into the medulla (70, 71) (Figure 1-2 C). Within the medulla, medullary thymic epithelial cells (mTECs) express a significant array of tissue restricted antigens that they may encounter in the periphery, such as insulin, by virtue of their expression of the transcription factor AIRE (72, 73). Here, should these developing T cells get a strong TCR signal (where strong is determined by high TCR:pMHC affinity/avidity), they will die via negative selection: thus 'central tolerance'.

Cells remain in the medulla for approximately 2-4 days, and those that do not encounter a strong TCR signal (hence do not robustly recognize "self") downregulate CD69 (Stage 5, Figure 1-2 C), which allows them to respond effectively to S1P signals in blood, and leave the thymus via the thymic vasculature as mature CD4 or CD8 SP cells (74) (Figure 1-2 C). Failures in central tolerance can lead to severe autoimmunity, as evidenced by humans and mice that lack AIRE (75, 76), mice that lack CCR7 and therefore have difficulty enter the thymic medulla (70), and mice that have defects in TCR signaling (discussed below) (77-79).

The balance between positive selection and negative selection is a conceptually fine line. On one hand, T cells are required to signal on self-peptide:MHC, but on the other,

if they signal too strongly, they are deleted. These outcomes are intrinsically linked to the strength of a TCR signal, and in settings where TCR signaling is impaired, selection fails and more autoreactive cells can escape (77-79). Developing thymocytes can only progress if they bear intermediate affinities to self-peptide, known as the “Goldilocks Zone” of TCR sensitivity. One mechanism that has been proposed to resolve the tension between positive and negative selection both occurring on self-peptide, is the concept of differential TCR “tuning” (80-87). Tuning the threshold for TCR sensitivity higher in the thymus would allow for more successful positive selection and more stringent negative selection, because self-peptides would be able to generate a stronger signal at both selection checkpoints (80-87). TCR detuning in the periphery then would limit autoreactivity because weak self-peptide signaling for positive selection is no longer needed, while stronger self-peptide signaling (i.e., autoimmunity) would be attenuated by detuning (80-87). How this mechanism is regulated remains unclear. Such a system should be tightly regulated to stages of development within the thymus. Chapter 3 of this thesis proposes that the tightly controlled expression of RAG1 (and thus RAG2) acts as a biological clock to time upregulation of signaling molecules during T cell development in order to enhance positive and negative selection.

1.6 RAG-mediated transcriptional alterations in development

When DNA breaks are detected, a transcriptional program is initiated by the DNA repair factor, ATM (88-90). Part of this transcriptional response is mediated via p53 activation, and drives a pro-apoptotic transcriptional program (91). Since DNA DSBs are a requirement of VDJ recombination, such a response would be detrimental to the development of T and B lymphocytes. More recently, two papers from the Sleckman group

have described the activation of RAG-specific transcriptional responses (92, 93). RAG-mediated DNA double stranded breaks in developing B cells activate the NFκB pathway, and induce the expression of pro-survival genes, other lymphocyte-specific genes, and a transcriptional repressor, SpiC, in order to repress expression of cell-cycle genes (92-94). One such pro-survival gene upregulated in this manner is Pim2 which acts as a survival factor by phosphorylating and deactivating the pro-apoptotic molecule BAD (95). Recently, the Bassing Lab has shown that in addition to transcriptional regulation of proteins such as Pim2, RAG1 can upregulate the expression of such proteins at the post-transcriptional level, and this is dependent on the N-terminal region of RAG1 (Unpublished Data, Personal Communication). Chapter 3 of this thesis proposes that the N-terminal region of RAG1, with a special focus on the E3 ligase domain of RAG1, controls the level of signaling molecules within T cells during development, to tune up TCR signaling.

1.7 Alternative translation, and the proteome

Absence of RAG1 leads to T/B SCID, however hypomorphic mutations in RAG1 that decrease but do not ablate activity can lead to Omenn Syndrome (44, 96, 97). Omenn Syndrome is a disease of concurrent autoimmunity and immunodeficiency. Typically, Omenn Syndrome patients present with B cell/T cell lymphopenia, autoreactive T cells, high IgE, eosinophilia, and erythroderma (98). Of particular note were a series of patients with frameshift mutations in the N-terminal region of RAG1 (99). Such mutations are predicted to lead to early protein translation termination and nonsense-mediated mRNA decay, yet the presence of T and B cells necessitated residual RAG activity, and hence presence of RAG1. It was hypothesized that internal translation initiation downstream of

the canonical translation initiation site (TIS) was able to lead to the production of N-truncated RAG1 isoforms (99).

Ribosome profiling implies internal translation initiation generates N-truncated isoforms for ~15% of the ~20,000 human proteins (100-103). Yet, our understanding of mechanisms that control internal translation initiation and how resulting N-truncated isoforms function normally and contribute to disease are at infancy (102). Protein synthesis typically starts at 5'AUGs in Kozak sequences (i.e. the TIS) (103-105). Translation initiation factors recruit ribosomes to mRNA 5' caps from where ribosomes scan 5'-to-3' and start making protein at the most 5' TIS encountered (103-105). It has been experimentally validated that ~30 mammalian mRNAs have internal ribosome entry sequences (IRES) that allow ribosomes to begin scanning downstream of the most 5' TIS and initiate translation at an internal TIS (105-112). Ribosomes also have been shown to pass the *Pim2* mRNA 5'AUG (leaky scanning) (113) or be halted by 5' hairpin structures and transferred to internal AUGs of the *Hsp70* mRNA (shunting) (114).

While the description of the human patients with frameshift mutations in RAG1 hypothesized that internal translation initiation lead to formation of internally translated RAG1 isoforms, the mechanism and whether these products are formed in normal circumstances remained unexplored. This is further addressed in Chapter 2. Additionally, the role of internal translation in human disease, and whether nonsense and frameshift mutations such as in RAG1 are pathogenic in other disease systems is unclear. This is an area requiring immediate attention.

1.8: Structure of thesis

While the structure and function of RAG1 in terms of its role in VDJ recombination has been well elucidated, in recent years it has been shown to have roles in addition to this recombination activity. Specifically, the N-terminal domain of RAG1 is necessary for the expression of a number of proteins during T cell development. This thesis addresses two pertinent points with regard to this:

In **Chapter 2**, I describe a mouse with an N-terminal nonsense mutation in RAG1 that leads to the formation of N-truncated RAG1. Through analysis of this mouse, and *in vitro* biochemical assays I describe how N-truncated isoforms are made via internal translation initiation. Most importantly, these alternative isoforms are present in normal settings as well. Since N-truncated RAG1 isoforms have differing abilities in terms of recombination activity, and that these isoforms can form part of the RAG1:RAG2 hetero-tetramer, I propose the novel concept that RAG1 normally exists in multiple isoforms and that these isoforms can have regulatory roles with regards to VDJ recombination and T/B cell development.

In **Chapter 3** I describe specifically how the E3-ligase domain is required for efficient expression of signaling molecules during T cell development, and that in the absence of functional RAG1 E3 ligase activity, positive and negative selection of T cells is disrupted. T cells that leave the thymus of these mice are hyper-functional and are able to cause more severe immune-driven pathology in a model of T cell transfer. That RAG1 is involved in T cell development in such a way is a novel conceptual advance that possibly links the ideas of suboptimal RAG function, and the development of autoimmunity, with a hypothesized mechanism.

Finally, in addition to discussion around the potential activities of RAG1 N-terminal truncated isoforms in normal biology, mechanisms of their formation, and future experiments to address such questions, in **Chapter 4** I also introduce the identification of another potentially disease-causing mutation (in NLRP1), that may lead to similar internal translation initiation. Mutations like these are often overlooked in whole exome/whole genome sequencing (WES/WGS) analyses, and examples like these should spur physicians to be cognizant of the potential effects of such mutations. *Enjoy.*

1.9: Figures

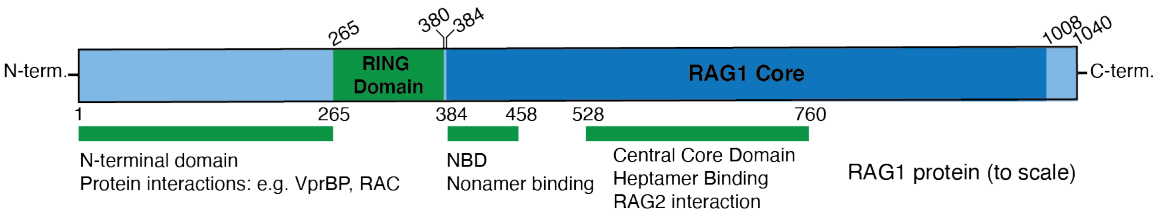


Figure 1-1: Cartoon of RAG1 structure and relevant domains
Modified from (29). Murine RAG1 consists of 1040 amino acids. Amino acids (AA) 384-1008 make up the RAG1-Core domain which is the minimally essential region required for recombination, and contains the nonamer and heptamer domains, and RAG2 binding domain. Within the N-terminal, non-core domain is the RING/E3-ligase domain (AA265-380), and regions that bind a number of additional proteins.

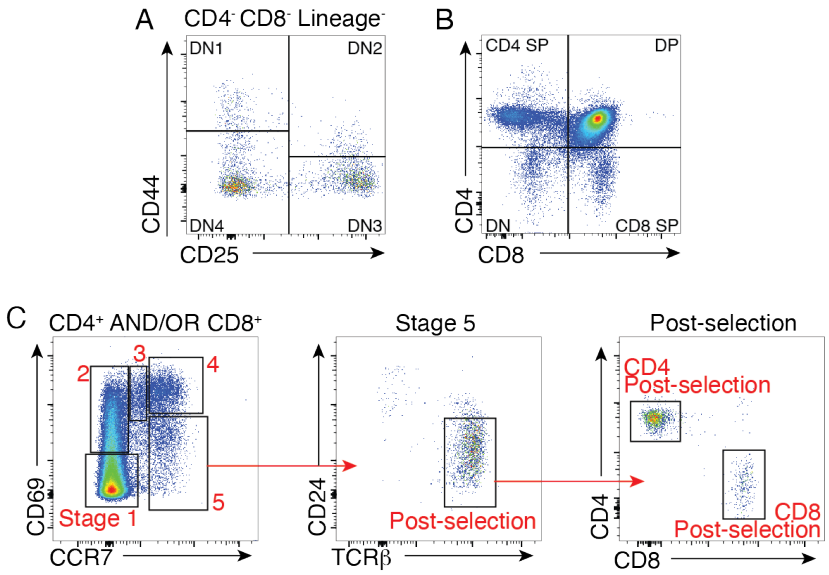


Figure 1-2: Flow cytometry gating of T cell developmental stages in the thymus

(A) Double negative (DN) cells showing DN1→DN2→DN3→DN4 progression. N.B. cKIT can be substituted for CD44.

(B) CD4 vs CD8 staining showing DN, DP (double positive), CD4 SP (CD4 Single Positive), and CD8 SP (CD8 single positive) populations.

(C) CCR7 and CD69 staining of non-DN populations to highlight the progression of cells through positive and negative development. The majority of cells sit in Stage 1, where TCRα is rearranged. Upon successful rearrangement of an in-frame TCRα that can pair with the previously rearranged and tested TCRβ, Stage 1 cells get a TCR signal that induces activation and expression of CD69, and then CCR7. Stage 3-4 cells that express CCR7 can migrate to the medulla of tha thymus were the majority of negative selection is undertaken. Upon downregulate of CD69, post-selected, CD4 and CD8 SP cells can migrate towards S1P and into blood vessels to leave the thymus (67, 68).

CHAPTER 2: ALTERNATIVE TRANSLATION OF RAG1 CREATES N-TERMINALLY TRUNCATED RAG1 ISOFORMS

2.1: Abstract

The RAG1 and RAG2 proteins are essential for the assembly of antigen receptor genes in the process known as VDJ recombination, allowing for an immense diversity of lymphocyte antigen receptors. Congruent with their importance, RAG1 and RAG2 have been a focus of intense study for decades. RAG1 has been studied as a single isoform, however, identification of a spontaneous nonsense mutation in the 5' end of the mouse Rag1 gene lead us to discover N-truncated RAG1 isoforms made from internal translation initiation. Mice homozygous for the mutation express smaller RAG1 isoforms, have defects in antigen receptor rearrangement, make fewer T and B cells, and present with some features of autoimmunity, similar to human Omenn Syndrome patients with 5' *RAG1* frameshift mutations. Given the seemingly inactivating *Rag1* mutation, it is striking that homozygous mutant mice do not have the expected severe combined immunodeficiency. We propose the possibility that evolution has garnered RAG1 and other important genes with the ability to form truncated proteins via internal translation to minimize the deleterious effects of 5' nonsense mutations. This mechanism also becomes intriguing as we consider the likely effects of nonsense mutations globally in the interpretation of whole genome sequencing.

2.2: Introduction

The RAG1/RAG2 hetero-tetramer has been extensively studied in its role as the enzyme responsible for rearranging antigen receptor genes in B and T cells in the process known as VDJ recombination (39, 115). RAG1 contains the catalytic domain necessary for DNA cleavage and thus VDJ recombination. Patients with loss-of-function mutations in either RAG1 or RAG2 suffer from severe combined immunodeficiency (SCID) with complete loss of mature T and B cells. In contrast, RAG1/2 mutations that reduce but do not ablate recombination activity lead to Omenn Syndrome (OS), an immunodeficiency of low B- and T-cell numbers, concurrent with autoimmunity (44, 96). Notably, Santagata and colleagues described patients with homozygous frameshift mutations in the N-terminal region of RAG1 that result in early translation termination and should not yield functional RAG1 protein (99). Despite this, these patients had low but detectable circulating B and T cells. RAG1 is translated from a single exon, thus alternative splicing was not considered. It was hypothesized that N-truncated RAG1 products arise from internal translation initiation at downstream AUG start sites, leading to a hypomorphic protein. Correlative with this hypothesis, expression of N-truncated cDNAs give rise to RAG1 proteins of smaller size that catalyze VDJ reactions in vitro. However, neither the mechanism by which these truncated RAG1 proteins are made nor whether they occur normally were elucidated (99).

During a routine analysis we identified a strain of mice in our colony with reduced T and B cells and increased activated T cell frequencies. Whole exome sequencing revealed a novel homozygous nonsense mutation at amino acid 60 (Q60X) which we termed the N-terminal stop or RAG1^{NX} mouse. This mutation, like frameshift mutations present in OS patients, would be predicted to cause premature translation termination,

and absence of functional RAG1. However, the presence of T and B cells in these mice, although at lower numbers, necessitated the presence of residual RAG1 activity, and western blot revealed the presence of smaller RAG1 isoforms in homozygous $RAG1^{NX}$ thymocytes. Here we show that the Q60X mutation in $RAG1^{NX}$ thymocytes results in expression of smaller RAG1 isoforms generated using internal translation initiation sites. Moreover, these smaller RAG1 isoforms are present in homozygous $RAG1^{WT}$ thymocytes. Our data demonstrate that an underappreciated mechanism of internal translation leads to multiple RAG1 isoforms in WT mice and allows for the escape from early truncating mutations in the gene that would otherwise have devastating immune consequences. There is significant evidence that the N-terminal regions of RAG1 have important roles in VDJ recombination (48, 49, 52, 54, 55), and RAG1 has been studied as a single isoform, containing the complete N-terminal region for over 30 years. The conceptual advance that multiple N-truncated RAG1 isoforms exist, even in wild type cells, is important as these N-truncated proteins demonstrate altered function and have not previously been interrogated as part of RAG biology. This novel murine model of OS provides an important platform to determine how N-truncated RAG1 proteins function normally and in the absence of full-length RAG1 protein.

2.3: Results

Novel nonsense mutation identified in 5' region of Rag1 gene

Routine analysis of peripheral blood mononuclear cells in a line of C57BL/6 mice within our colony revealed B and T cell lymphopenia of unknown cause (data not shown). Whole exome sequencing revealed a homozygous nonsense mutation early in the *Rag1* gene at nucleotide position 187 (c187C>T), which was confirmed by Sanger sequencing (Figure

2-1A/B). This mutation introduces a stop codon that is predicted to terminate translation at amino acid 60 of RAG1 (Q60X) and result in no functional RAG1. We named this the $RAG1^{NX}$ mouse. However, unlike complete RAG1 deficiency, this homozygous mutation does not result in complete absence of T and B cells. We therefore checked for residual RAG1 protein by western blot on bulk homozygous $RAG1^{NX}$ thymocytes and show a number of proteins of smaller sizes than full-length (FL), wild-type RAG1 (Figure 2-1C). Truncating frameshift mutations in the 5' region of the Rag1 gene in human patients are unexpectedly associated with Omenn Syndrome (OS) rather than SCID (99). Thus, we considered that the $RAG1^{NX}$ mouse may represent a model of OS to investigate how such mutations retain enough RAG activity to prevent complete SCID.

$RAG1^{NX}$ mice have lymphocyte developmental blocks at antigen receptor rearrangement steps

Given the mutation in RAG1, we sought to determine the effect on T and B cell development. $RAG1^{NX}$ thymuses demonstrate a significant reduction in total cellularity, accounted for by a ~3-fold reduction in the number of double positive ($CD4^+CD8^+$) and mature single positive cells (Figure 2-2 A/B). The major developmental block occurs at the DN3 to DN4 transition (Figure 2-2 C/D). We find an analogous block in B cell development, at the pro-B to pre-B transition (Figure 2-2 E/F). These phenotypes are inherited in an autosomal recessive manner, as $RAG1^{NX}$ heterozygotes are phenotypically indistinguishable from homozygous $Rag1^{WT}$ mice.

The DN3-DN4 transition for developing T cells, and pro-B to pre-B transition for B cells requires the productive rearrangement of the TCRβ allele, or immunoglobulin heavy

(IgH) chain for T and B cells respectively, and subsequent signaling through either the pre-TCR or pre-BCR (116). Blocks at these developmental stages can be due to defective recombination efficiency, therefore we sought to determine whether RAG1^{NX} cells had decreased VDJ recombination.

RAG1^{NX} thymocytes have reduced TCRVβ recombination and an altered Vβ TCR repertoire

To examine recombination efficiency and Vβ repertoire in RAG1^{NX} mice, we sorted DN3 thymocytes and assayed Vβ-DβJβ rearrangements by TaqMan PCR. We find substantially lower than normal levels of rearrangements of all Vβ gene segments in RAG1^{NX} DN3 cells (Figure 2-3 A/B). We also examined the Vβ repertoire of total thymocytes as an indication of global Vβ repertoire differences and see decreases in the usage of many Vβ alleles, but no differences in others suggesting global repertoire has shifted (Figure 2-3 C/D).

Next we determined the Vβ repertoire of thymocytes that produce a functional TCRβ by surface staining and show that there is an altered repertoire at both DP (Figure 2-4 A/C) and SP stages (Figure 2-4 B/D). This supports the global repertoire analysis by PCR (Figure 2-3 C/D). By analyzing the ratio of Vβ usage at the DP versus the SP stage we show that there is altered selection of certain Vβ proteins, such as less selection of Vβ4, and greater selection of Vβ5 (Figure 2-4 E).

RAG1^{NX} mice have reduced numbers of mature B and T cells and increased frequencies of Tregs and Tmemory cells

RAG1^{NX} mice have an approximate 50% reduction in peripheral T and B cell counts. This phenotype is inherited in an autosomal recessive manner, as RAG1^{NX} heterozygotes are phenotypically indistinguishable from homozygous Rag1^{WT} mice (Figure 2-5 A). There is no effect on dendritic cells, neutrophils, and inflammatory monocytes, but a slight increase in the numbers of splenic NK cells (Figure 2-5 A). RAG1^{NX} mice had higher frequencies of Tregs (Figure 2-5 B), and memory phenotype CD4^{+} and CD8^{+} T cells (Figure 2-5 C/D). This phenotype is consistent with the leaky-SCID/OS phenotype observed in humans and other mouse models of OS (96, 97, 117, 118).

RAG1^{NX} mice develop increased age-associated B cells and anti-nuclear antibodies

Given the similarity of the Q60X mutation present in RAG1^{NX} mice to lesions found in a number of OS patients, we assessed phenotypes typically associated with development of autoimmunity (99). *Ex vivo* stimulated CD4^{+} T cells from young (4-6week old) RAG1^{NX} mice show a heightened propensity to express IFNγ and IL-17A, and CD8^{+} T cells express more IFNγ and TNFα than WT counterparts (Figure 2-6 A-D). OS patients often present with erythroderma, eosinophilia, high circulating IgE levels, and Th2 skewing including T cell IL-4 and IL-5 production. While it is clear that young RAG1^{NX} mice preferentially adopt a Th1 skewed T cell phenotype, more characteristic of the C57BL/6 strain (119), we were interested to see whether canonical OS phenotypes arose as the mice aged. RAG1^{NX} mice aged 20-22-months have no difference in immune cell populations and T and B cell counts have equilibrated in mice by this age (Figure 2-6 E). There is no evidence of

increased eosinophils. Overall RAG1^{NX} mice have no overt autoimmune pathology (data not shown) but feature increased CD11b^{+}CD11c^{+} age-related B cells (ABCs) (Figure 2-6 F), that display increased Tbet expression (Figure 2-6 G). ABCs are often associated with a predisposition towards anti-nuclear antibody (ANA) positive autoimmune diseases (120). Accordingly, aged RAG1^{NX} mice have higher ANA levels (Figure 2-6 H/I). Overall these data suggest that while RAG1^{NX} mice do not have overt OS disease as seen in humans, they may have a predisposition towards autoimmunity.

Smaller RAG1 isoforms arise via internal translation initiation

Given the position of the RAG1^{NX} nonsense mutation (Figure 2-1 A), and the presence of smaller than normal RAG1 proteins in RAG1^{NX} cells (Figure 2-1 C), we were intrigued as to how these smaller proteins form. The description of OS patients with 5' RAG1 frameshift mutations hypothesized that alternate N-truncated isoforms arise via translation from alternative translation initiation sites (99). While this report showed that N-terminally truncated RAG1 proteins function in recombination assays, whether alternative translation initiation creates these isoforms was not directly assessed. Interestingly, even RAG1^{WT} thymocytes have expression of smaller proteins (Figure 2-1 C). These proteins have been noted and designated previously as 'break-down products' (121). We hypothesized that these smaller proteins are naturally occurring isoforms created from translation of alternative translation initiation sites (TIS's). To test this, we created RAG1 proteins epitope-tagged at N-terminal (FLAG) and C-terminal (HA) ends. Probing for FLAG only detects full-length RAG1 translated from the canonical TIS while probing for HA would reveal all possible internally translated isoforms (Figure 2-7 A). Site-directed mutagenesis was performed to introduce the Q60X mutation and to mutate the canonical initiating

methionine (iM) codon, AUG, to isoleucine (AUC, iMI). RAG1-WT, RAG1-Q60X, and RAG1-iMI constructs were transfected into 293T cells and FLAG and HA expression was analyzed. FLAG expression is only detected in the WT transfected cells (Figure 2-7 B), whereas the C-terminal HA-tag is detected in all RAG1-transfected cells (Figure 2-7 C). The full-length band at ~120kDa is detected by the FLAG-tag in WT by western blot, whereas Q60X and iM1I mutants are absent (Figure 2-7 D). Analysis of the C-terminal HA tag reveals a number of proteins in all cases. The Q60X and iM1I mutant RAG1 constructs could only produce the smaller isoforms (Figure 2-7 E). To interrogate whether the smaller bands were indeed produced by internal translation, the first 5 internal methionines (AUG) were mutated to isoleucines (AUC) in a sequential and additive manner (i.e. M3I = M3I+M2I+M1I). As before, RAG1 constructs were transfected into 293T cells and only the full-length, ~120kDa RAG1 isoform is detected via FLAG probing (Figure 2-7 F). As hypothesized, the potential translation initiation sites at M1, M4, and M5 are required for the expression of their respective smaller isoforms as observed by the loss of specific bands upon AUG to AUC mutation (Figure 2-7 G). Whether the M2 TIS gives rise to the band at ~80kDa is inconclusive, as M2I mutation shows a slight decrease in protein quantity at the putative M2 band, but not complete ablation as with M1, M4, and M5. It is clear that mutation of the TIS at M3 does not impact expression of the putative M3 isoform. We also interrogated the strength of the Kozak sequences surrounding the putative TIS sites using a published formula (122) and term these “Noderer scores”. We show that M1, M2, M4, and M5 all have extremely strong Kozak sequences (Figure 2-7 H). Comparison of the Noderer scores (122) of Kozak sequences surrounding the putative TIS sites with random, out of frame, AUGs within the RAG1 transcript highlights that these high scores are unlikely to be random, and may be selected for (Figure 2-7 I). We therefore sought to

determine whether these strong Kozak sequences surrounding internal translation initiation sites in RAG1 were conserved across species.

Strong internal RAG1 translation initiation sites are conserved across species

Evolutionary selection for these alternative TIS sites is supported by conservation of strong internal TISs across species (Figure 2-8). Given the degenerate nature of codon specificity, it may not be expected that these Kozak sequences would be selected for on the basis of the full-length protein alone, as synonymous mutations would not impact this, yet they may impact the natural formation of N-truncated RAG1 isoforms. Overall, these data highlight that RAG1 can be translated from internal TIS's, and that these N-truncated RAG1 isoforms may have been evolutionarily selected.

Smaller RAG1 isoforms can co-immunoprecipitate with RAG2

Finally, we were interested in whether the naturally occurring N-truncated RAG1 isoforms of WT RAG1 could co-immunoprecipitate with RAG2, since this would likely be essential if these smaller isoforms were to have important functions. RAG1-HA and RAG2-GFP constructs were co-transfected into 293T cells. RAG2-GFP was immunoprecipitated using anti-GFP beads and isolated protein was analyzed by SDS-PAGE and probed for RAG1-HA using an anti-HA antibody (Figure 2-9). Clearly, N-truncated RAG1 isoforms can by co-immunoprecipitated with RAG2 suggesting these isoforms can readily form part of the hetero-tetrameric RAG complex. This would be a fundamental shift in the understanding of normal RAG biology, and future investigation of this will be essential if we are to understand RAG biology fully.

2.4: Discussion

Here we describe for the first time the mechanism of formation of alternative RAG1 isoforms. These RAG1 isoforms arise as a result of internal translation initiation and lack N-terminal regions of RAG1. They were revealed upon identification of a novel nonsense mutation in RAG1, a genetic lesion that is very similar to that present in a number of Omenn Syndrome patients. We therefore present here a novel murine model of OS-like disease that both describes mechanistically how functional RAG1 protein can be made despite early nonsense or frameshift mutations, and a more detailed description of T and B cell development in mice with such mutations. This provides a model to further interrogate how autoimmunity manifests concurrently with immunodeficiency, and the effect that RAG1 protein or defective recombination has to play in this.

The Q60X nonsense mutation in RAG1 would upon initial observation be expected to result in a non-functional peptide that could not participate in VDJ recombination, as it should lack the catalytic domain of RAG1 (the $RAG1^{core}$ domain). However, some internal translation initiation products provide N-terminal truncated RAG1 proteins that contain the $RAG1^{core}$ domain. In fact, many of the phenotypes displayed by the $RAG1^{NX}$ mouse replicate that of the $RAG1^{core}$ mouse including blocks in development at antigen rearrangement stages (47, 48). $RAG1^{core}$ mice have defective rearrangement, and similarly the $RAG1^{NX}$ phenotype is likely due to inefficient recombination activity (47, 48). Interestingly, altered selection of $V\beta$ proteins from the DP to SP stage is also seen in the $RAG1^{core}$ mice (48). This altered $V\beta$ usage is independent of choice of $TCR\beta$ rearrangement, but may be due to altered $TCR\alpha$ repertoire, or reduced secondary $V\alpha$-to-$J\alpha$ rearrangements. Thus, both in terms of lymphocyte numbers and VDJ rearrangement,

the RAG1^{NX} mouse phenocopies RAG1^{core}, consistent with the RAG1^{NX} mutation unexpectedly generating smaller N-truncated isoforms.

As a result of decreased recombination efficiency, young RAG1^{NX} mice expectedly display peripheral B and T cell lymphopenia. The increased frequencies of memory phenotype T cells is likely a result of homeostatic expansion of T cells within lymphopenic hosts (123), and increased Treg frequencies because Tregs outcompete conventional T cells (124) or convert from conventional T cells in a lymphopenic environment (125). This homeostatic expansion is sufficient to fill the available T/B lymphocyte niche upon aging of the mice. The role of lymphopenia in driving autoimmune disease has been postulated a number of times, yet it is difficult to tease apart the relative contributions of lymphopenia from altered lymphocyte development, defective VDJ rearrangement, and altered antigen receptor repertoires. It is attractive to speculate that homeostatic expansion of few T cells drives oligoclonality, and that those T cells that escape development with higher self-reactivity can expand disproportionally.

While the RAG1^{NX} mouse has a similar genetic lesion to humans with OS, and similar lymphopenia, the pathology experienced by OS humans is not entirely replicated. However, there is some evidence of immune dysregulation in that RAG1^{NX} T cells can more efficiently make effector cytokines *ex vivo*, and aged RAG1^{NX} mice have an increase in age-associated B cells and increased ANAs. It is also important to note that, in comparing the RAG1^{NX} mouse to two other murine models of OS, the MM mouse (117), and the R229Q RAG2 mouse (118), we see a number of similarities and differences. All three models display a block in development of B and T cells at antigen receptor gene rearrangement steps, increases in memory phenotype T cells, and altered TCR

repertoires. The MM and R229Q RAG2 mice show some facets of canonical OS, such as erythroderma and high IgE, but differ substantially in other aspects suggesting the specific genetic lesion, mouse strain (which differs between all three mice), and environment in which they are housed all may contribute to the observed autoimmune phenotypes. Of note, the C57BL/6 strain of mice that our $RAG1^{NX}$ mutation is present in, is notoriously resistant to autoimmunity in the absence of additional environmental stimuli (126). Similarly, the R229Q RAG2 mutation on the C57BL/6 background does not present with overt autoimmunity (127).

Finally, we describe the mechanism of formation of alternative RAG1 isoforms, and while this was revealed in a setting where full-length RAG1 could not be made, these appear to exist in WT mice as well. Additionally, strong internal translation initiation sites appear to be conserved across species, and N-truncated RAG1 isoforms are able to be co-immunoprecipitated with RAG2. Should these N-truncated RAG1 isoforms exist naturally in biologically relevant amounts then this would be a fundamental shift in the understanding we have of RAG biology. It is feasible to postulate that since N-truncated RAG1 has reduced recombination efficiency, then the normal presence of truncated RAG1 isoforms may act as important negative regulators of RAG efficiency. In fact, inefficient VDJ recombination is hypothesized as a mechanism to prevent the expression of multiple rearranged TCR or BCR alleles from a single T or B cell in the process of allelic exclusion (128, 129). Additionally, as will be described in detail in Chapter 3, the N-terminal region of RAG1 appears to have a growing number of roles in addition to its VDJ activity. Naturally occurring N-truncated isoforms would therefore be lacking these capabilities as well. These ideas regarding the role of internally translated RAG1 isoforms in lymphocyte

development, and what the identification of an ever-increasing number of internally translated protein isoforms could mean outside of RAG1 will be expanded upon in Chapter 4.

2.5: Figures

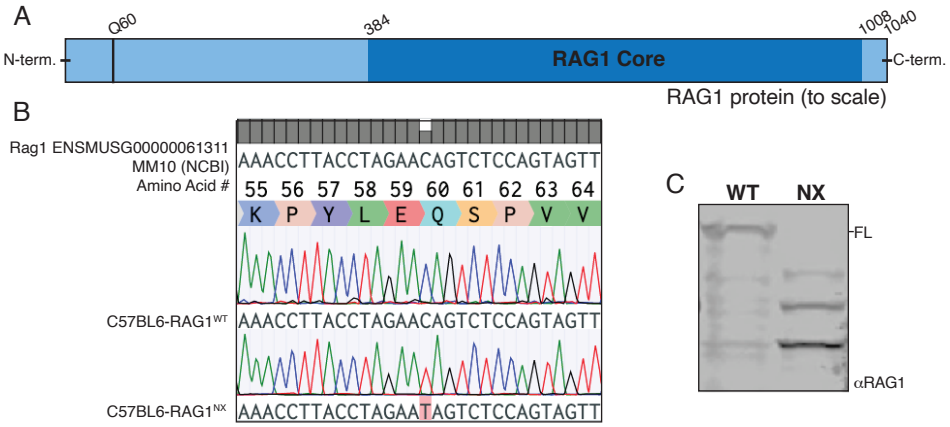


Figure 2-1: Novel nonsense mutation identified in 5' region of Rag1 gene
(A) Diagram of RAG1 protein showing relative position of novel nonsense mutation (Q60X). **(B)** Sanger sequencing tracks aligning sequences from $RAG1^{WT}$ and $RAG1^{NX}$ mice in our colony to that of the reference mm10 genome (NCBI). **(C)** RAG1 protein expressed in bulk thymocytes was analyzed by SDS-PAGE and Western blot.

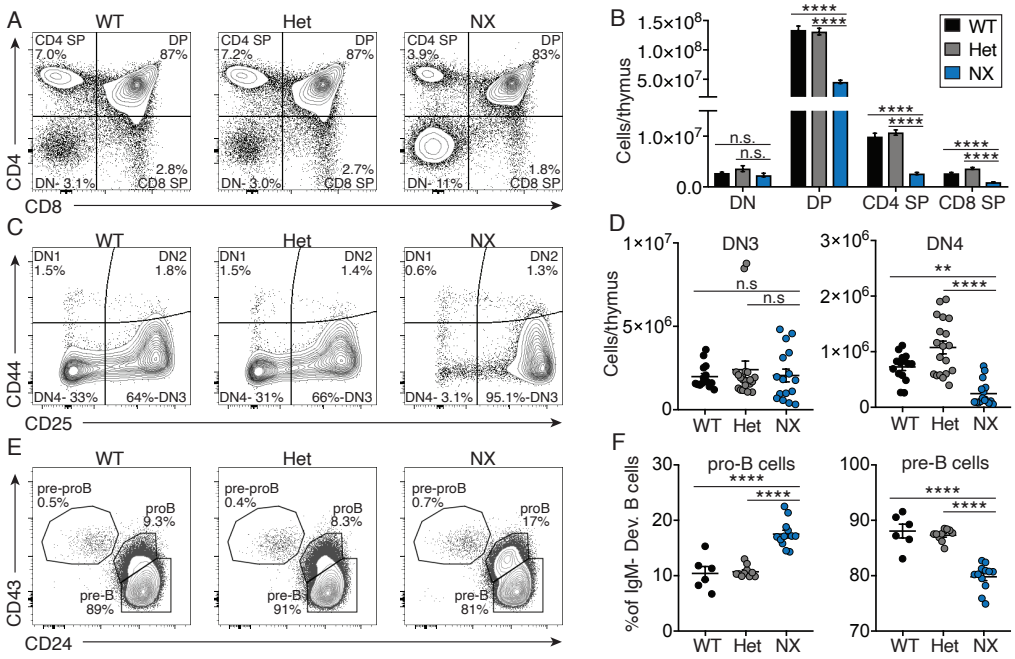


Figure 2-2: RAG1^{NX} mice have lymphocyte developmental blocks at antigen receptor rearrangement steps

Thymi from 4-5week old WT, NX, and Het littermates were analyzed by flow cytometry. **(A)** CD4 vs CD8 frequencies, and **(B)** total cell counts per thymus. Gating: Live, singlets, Dump$^-$(B220, CD11b, CD11c, Gr1, NK1.1, Ter119), TCR$\gamma\delta^-$. **(C)** DN1-4 frequencies, and **(D)** total cell counts per thymus. Gating: CD4$^-$,CD8$^-$,TCRβ^{lo}. Early B cell development was analyzed in the bone marrow. **(E,F)** Frequencies of early B cell progenitors of B220$^+$ CD93$^+$IgM$^-$ cells were quantified. Pre-gating: Live, singlets, dump$^-$ (TCRβ, NK1.1, Ter119, CD11c, Gr1). Data combined from at least 3 independent experiments. Bars indicate mean +/- SEM. Statistics: 1-way ANOVA with Tukey-HSD post-test.

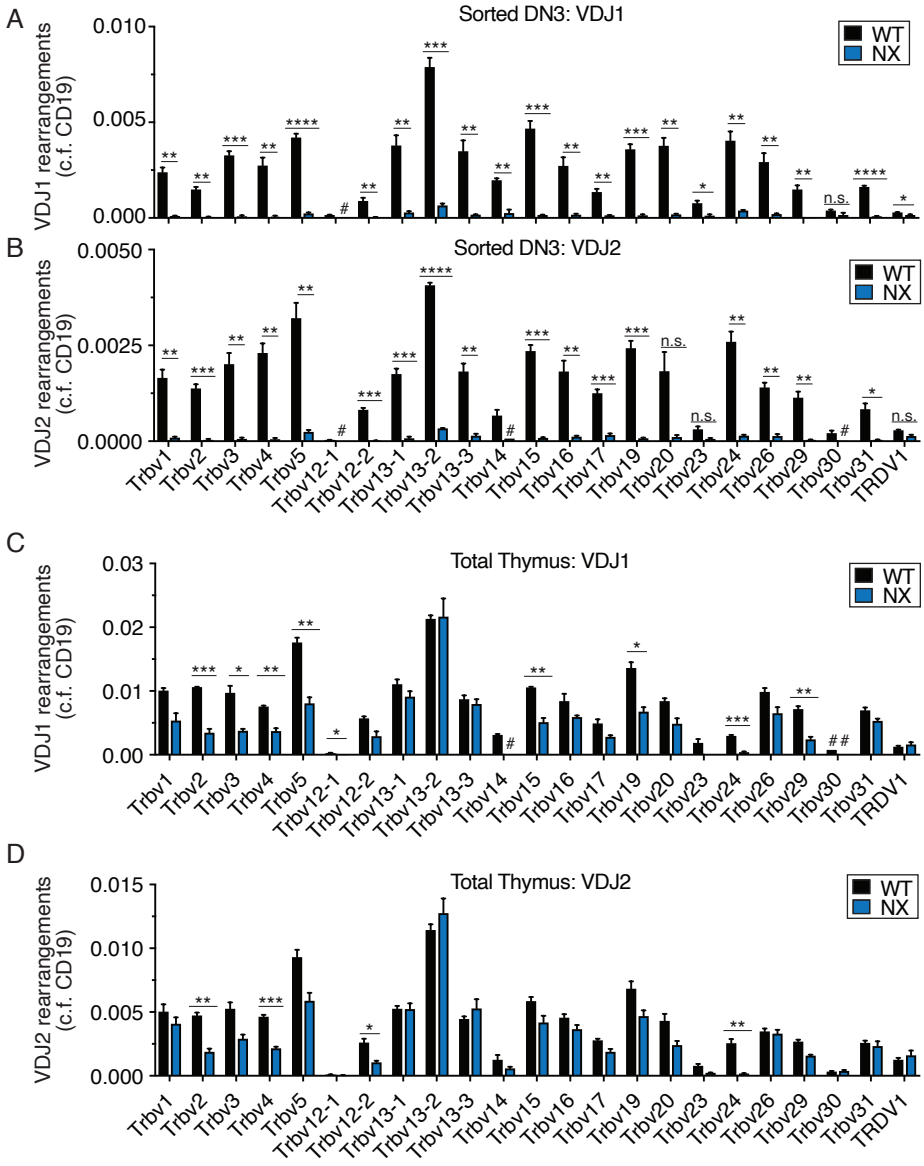


Figure 2-3: RAG1^{NX} thymocytes have reduced TCRVβ recombination and an altered Vβ TCR repertoire

Sorted DN3 thymocytes from WT and NX mice were assayed by qPCR for frequencies of respective Vβ to DβJβ1 rearrangements **(A)**, and Vβ to DβJβ1 rearrangements **(B)**, 4 mice per genotype. #not detected. Gating: live, singlets, dump$^-$ (B220, CD8, CD4, CD11b, CD11c, Gr1, NK1.1, Ter119), CD4$^-$, CD8$^-$, CD25$^+$, CD44$^-$. Vβ to DβJβ1 rearrangements **(C)**, and Vβ to DβJβ2 rearrangements **(D)** were also completed on bulk, unsorted, thymocytes as an indication of global Vβ usage.

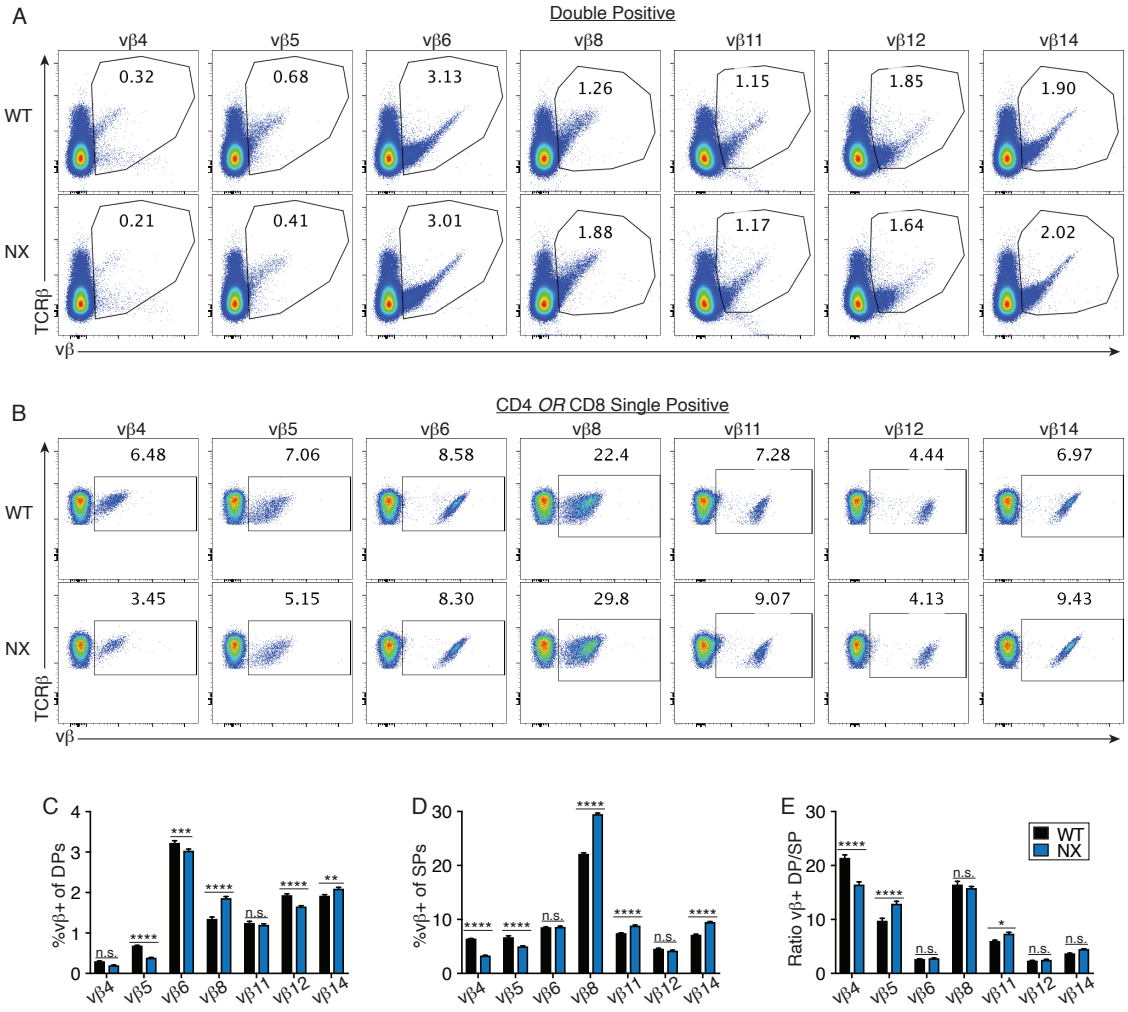


Figure 2-4: RAG1^{NX} thymocytes utilize different TCRvβ alleles and have altered selection of alleles

TCRvβ repertoire was assayed by flow cytometry of **(A)** DP thymocytes, and **(B)** SP thymocytes. Summary data presented for DP thymocytes and SP thymocytes in **(C)** and **(D)** respectively. Gating: Live singlets, DP=CD4$^+$CD8$^+$, SP=CD4$^+$ *OR* CD8$^+$. **(E)** The ratio of vβ usage by DP vs. SP thymocytes was calculated. Data combined from 2-independent experiments. Error bars indicate SEM. Statistics: multiple t-tests with Holm-Sidak correction.

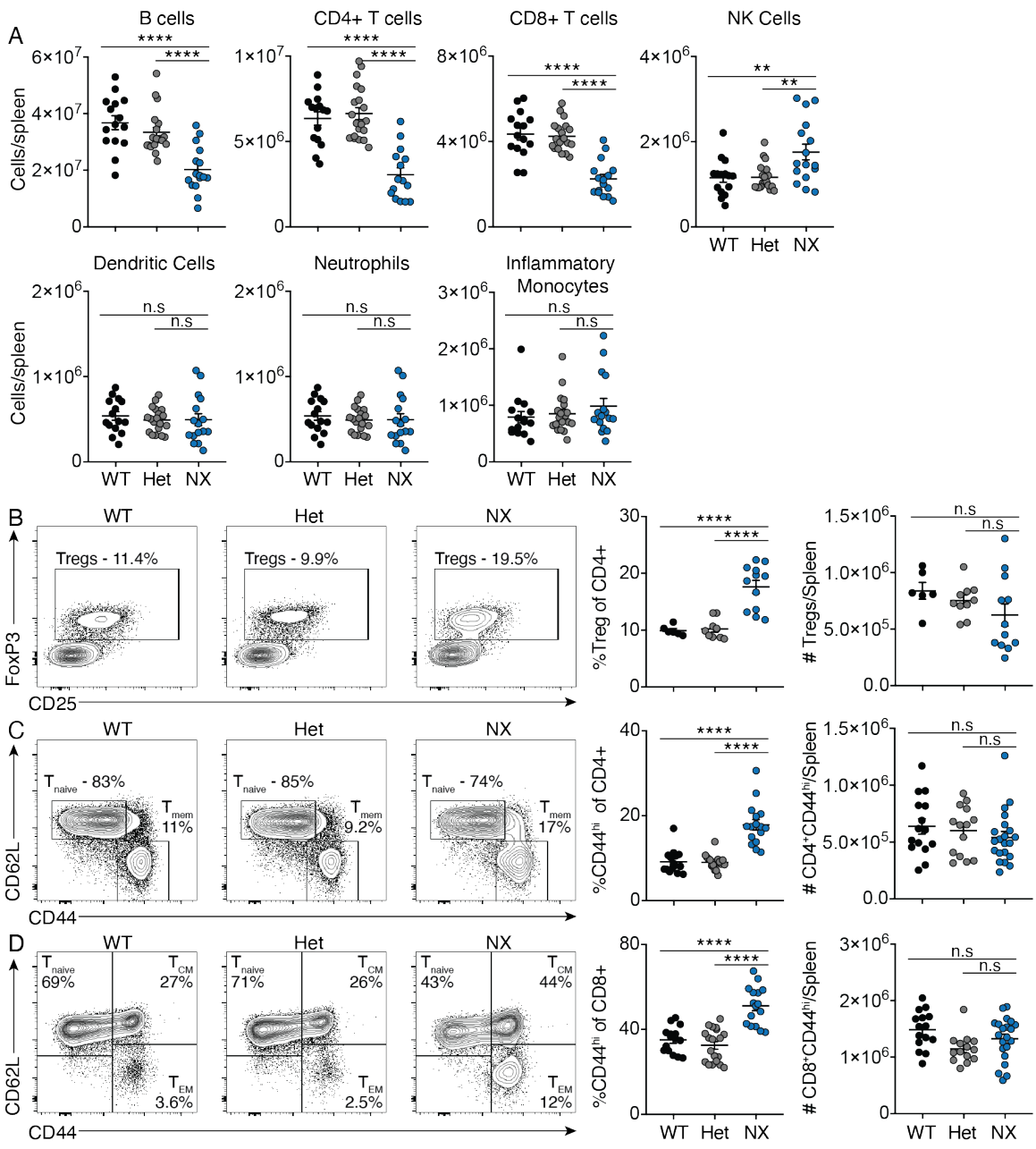


Figure 2-5: RAG1^{NX} mice have reduced numbers of mature B and T cells and increased frequencies of Tregs and Tmemory cells

(A) Numbers of leukocytes were enumerated from spleens of 4-5week old RAG1^{WT}(WT), RAG1^{NX}(NX) or Heterozygous (Het) littermates by flow cytometry. Relative frequencies and numbers of peripheral Tregs (B), memory CD4^{+} T-cells (C), and memory CD8^{+} T-cells (D) in the spleen were quantified. Data combined from at least 3 independent experiments. Bars indicate mean +/- SEM. Statistics: 1-way ANOVA with Tukey-HSD post-test. Gating: Live; singlets; B cells (CD19^{+}, TCRβ^{-}); T Cells (TCRβ^{+}, CD19^{-}, CD4^{+} or CD8^{+}); NK cells (TCRβ^{-}, CD19^{-}, NK1.1^{+}); Dendritic

cells (TCRβ^-, CD19$^-$, NK1.1$^-$, CD11c$^+$); Neutrophils (TCRβ^-,CD19$^-$,NK1.1$^-$,CD11c$^-$,CD11b$^+$,Ly6G$^+$); Inflammatory Monocytes (TCRβ^-, CD19$^-$, NK1.1$^-$, CD11c$^-$, CD11b$^+$, Ly6G$^-$, Ly6C^{hi}). T_{CM}=Central Memory, T_{EM}=Effector Memory.

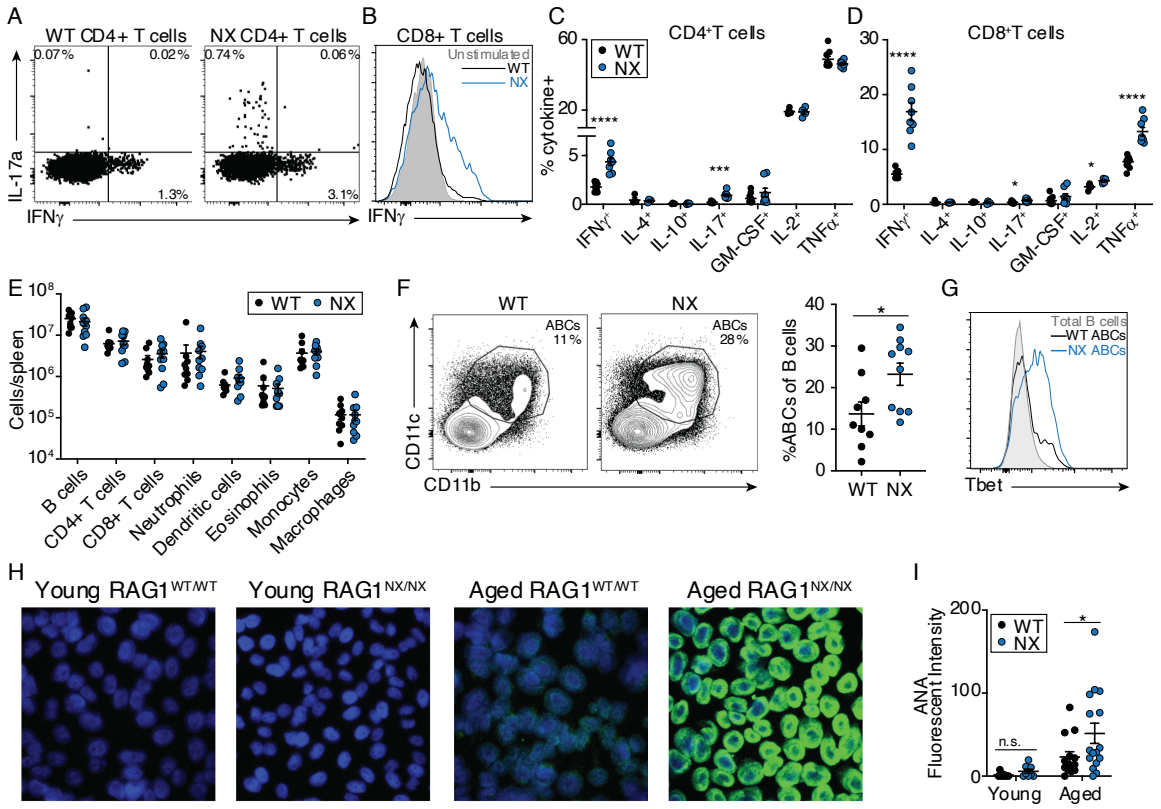
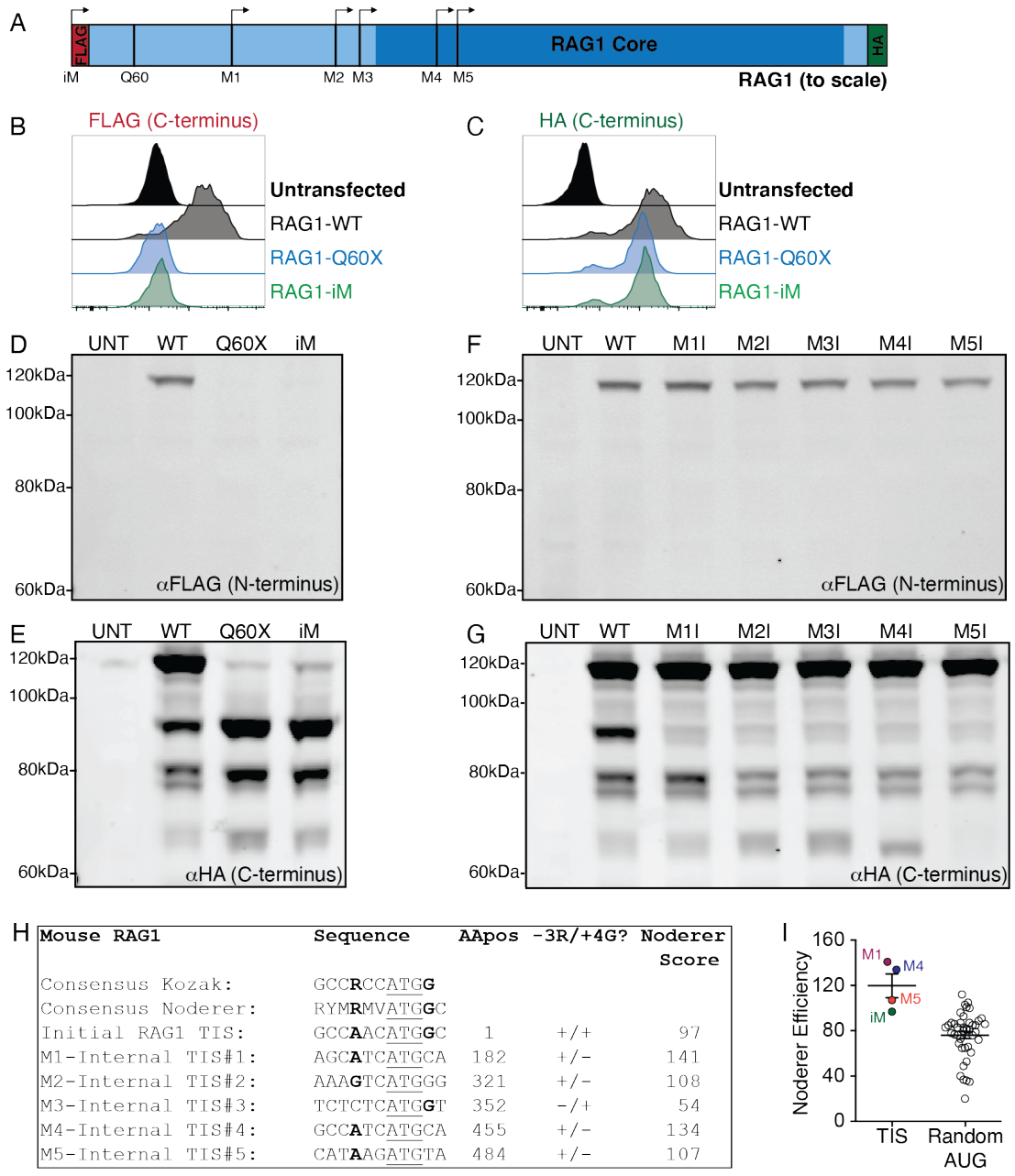


Figure 2-6: RAG1^{NX} mice have increased frequencies of cytokine producing T cells, and develop increased age-associated B cells and anti-nuclear antibodies

(A-D) T-cells from 4-5week old WT and NX mice were stimulated ex-vivo with PMA and ionomycin and cytokine expression was measured by flow cytometry. Statistics: 1-way ANOVA with Tukey-HSD post-test. Gating: Live singlets, CD90.2$^+$, CD4$^+$ *or* CD8$^+$. **(E)** 20-22month old WT and NX mice were compared for peripheral leukocyte populations. Gating: Neutrophils(CD3$^-$,Ly6G$^+$), B cells(CD3$^-$, Ly6G$^-$, CD19$^+$), T cells (CD3$^+$, CD19$^-$, Ly6G$^-$, CD4$^+$ *OR* CD8$^+$), Eosinophils (CD3$^-$, CD19$^-$, Ly6G$^-$, Siglec-F$^+$), Macrophages (CD3$^-$, CD19$^-$, Ly6G$^-$, Siglec-F$^-$, F4/80$^+$), Monocytes (CD3$^-$, CD19$^-$, Ly6G$^-$, Siglec-F$^-$, F4/80$^-$, CD11b$^+$). **(F)** Frequency of age-related B cells (ABCs) of total B cells was quantified in 20-22month old female WT and NX mice. Statistics: Mann-Whitney test. **(G)** Tbet staining on WT ABCs vs. NX ABCs of ~20week old mice. **(H-I)** Antinuclear antibody (ANA) staining intensity on Hep-2a cells was assayed for in young (4-5week old) and old (20-22month old) WT and NX mice. Bars indicate mean +/- SEM. Statistics: 1-way non-parametric ANOVA with Tukey-HSD post-test.



Mouse RAG1	Sequence	AApos	-3R/+4G?	Noderer Score
Consensus Kozak:	GCC**R**CCATG**G**			
Consensus Noderer:	RYM**R**MVATG**G**C			
Initial RAG1 TIS:	GCC**A**ACATG**G**C	1	+/+	97
M1-Internal TIS#1:	AGC**A**TCATGCA	182	+/-	141
M2-Internal TIS#2:	AAA**G**TCATGGG	321	+/-	108
M3-Internal TIS#3:	TCTCTCATG**G**T	352	-/+	54
M4-Internal TIS#4:	GCC**A**TCATGCA	455	+/-	134
M5-Internal TIS#5:	CAT**A**AGATGTA	484	+/-	107

Figure 2-7: Smaller RAG1 isoforms arise via internal translation initiation

(A) RAG1 schematic highlighting positions of FLAG and HA epitope tags, Q60X mutation, and internal methionines that serve as potential TISs. **(B/C)** Expression of RAG1-WT, Q60X, and iM constructs in 293T cells assayed for FLAG and HA expression by flow cytometry. Pre-gated on GFP+ cells. **(D/E)** Expression of RAG1-WT, Q60X, and iM mutants in 293T cells analyzed by SDS-PAGE and Western blot for FLAG **(D)** and HA **(E)** expression. **(F/G)** RAG1 WT vs. internal methionine mutants (M1-M5) expressed in 293T cells were analyzed by SDS-PAGE and Western blot for FLAG **(F)** and HA **(G)** expression. All blots are representative of at least 3 independent experiments. **(H)** Analysis of strength of internal putative TIS Kozak sequences compared to canonical Kozak sequence, and Noderer consensus sequence(122) AA Pos refers to the amino

acid position of the methionine in the RAG1 gene. -3R/+4G indicates the presence of a purine at the -3 position (relative to A of AUG = +1), and a G present at the +4 position. **(I)** Noderer scores of putative TIS's compared to Noderer scores surrounding random, out of frame AUG sequences within murine RAG1 transcript.

Initial TIS		AA Pos	-3R/+4G?	Noderer Efficiency
Consensus Kozak:	GCCRCCATGG			
Consensus Noderer:	RYMRMVATGGC			
Mus musculus(mouse)	**GCCAACATGGC**	1	+/+	97
Homo sapiens(human)	**GCCAGCATGGC**	1	+/+	107
Pan troglodytes(orangutan)	**GCCAGCATGGC**	1	+/+	107
Macaca mulata(rhesus macaque)	**GCCAGCATGGC**	1	+/+	107
Rattus Norvegicus(Rat)	**GCCAAAATGGC**	1	+/+	101
Sus scrofa(Wild boar)	**GCCAGCATGGC**	1	+/+	107
Mustelo putorius furo(ferret)	**GCCAGCATGGC**	1	+/+	107

M1 (AA~182)		AA Pos	-3R/+4G?	Noderer Efficiency
Consensus Kozak:	GCCRCCATGG			
Consensus Noderer:	RYMRMVATGGC			
Mus musculus(mouse)	**AGCATCATGCA**	182	+/-	141
Homo sapiens(human)	**AGCATCATGCA**	183	+/-	141
Pan troglodytes(orangutan)	**AGCATCATGCA**	183	+/-	141
Macaca mulata(rhesus macaque)	**AGCATCATGCA**	183	+/-	141
Rattus Norvegicus(Rat)	**AGCATTATGCA**	182	+/-	122
Sus scrofa(Wild boar)	**AGCTTCATGCA**	183	-/-	66
Mustelo putorius furo(ferret)	**CGCATCATGTA**	183	+/-	119

M4 (AA~455)		AA Pos	-3R/+4G?	Noderer Efficiency
Consensus Kozak:	GCCRCCATGG			
Consensus Noderer:	RYMRMVATGGC			
Mus musculus(mouse)	**GCCATCATGCA**	455	+/-	134
Homo sapiens(human)	**GCCATCATGCA**	458	+/-	134
Pan troglodytes(orangutan)	**GCCATCATGCA**	458	+/-	134
Macaca mulata(rhesus macaque)	**GCCATCATGCA**	458	+/-	134
Rattus Norvegicus(Rat)	**GCCATCATGCA**	455	+/-	134
Sus scrofa(Wild boar)	**GCCATCATGCG**	458	+/-	92
Mustelo putorius furo(ferret)	**GCCATCATGCA**	458	+/-	134

Figure 2-8: Strong internal translation sites are conserved across species
The strength of Kozak sequence was compared around the annotated TIS (iTIS) and internal M1 and M2 TIS' compared to related mammalian species as according to Noderer Score. AA Pos refers to the amino acid position of the methionine in the RAG1 gene. -3R/+4G indicates the presence of a purine at the -3 position (relative to A of AUG = +1), and a G present at the +4 position.

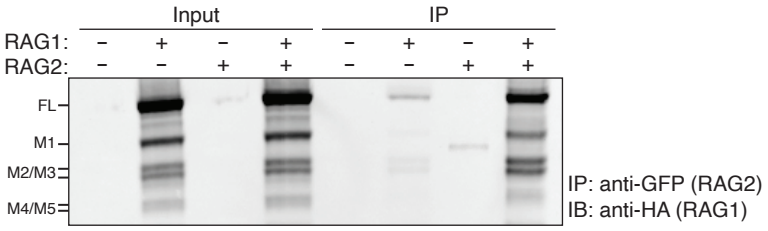


Figure 2-9: Smaller isoforms of RAG1 can co-immunoprecipitate with RAG2
RAG1 (HA tagged) and RAG2 (GFP tagged) were co-transfected into 293T cells. RAG2 was immunoprecipitated using anti-GFP beads. Protein was then subjected to SDS-PAGE and immunoblotted for HA to detect full-length and all N-truncated RAG1 isoforms. +/- indicates presence or absence of plasmid transfected into 293Ts. Input is cell lysate that has not been immunoprecipitated. IP = immunoprecipitation. IB = immunoblot. FL = full length RAG1, M1-M5 = likely protein products of internal translation from methionines 1-5 as indicated in Figure 2-7 A.

CHAPTER 3: RAG1 E3-LIGASE ACTIVITY IS REQUIRED FOR EFFICIENT POSITIVE AND NEGATIVE SELECTION OF DEVELOPING THYMOCYTES

3.1: Abstract

The formation of a diverse T cell receptor (TCR) repertoire is at odds with the need to prevent the development of autoreactive TCRs. Developing T cells are subjected to positive and negative selection in the thymus to ensure both TCR functionality, and to delete T cells bearing autoreactive TCRs. Both positive and negative selection signals are driven by self-peptide MHC interactions, which is seemingly incongruous, and requires developing T cells to have intermediate affinities to self-antigen. It has been proposed that thymocytes have heightened signaling during development such that T cells selected on self-peptide will not react to the same positive selecting peptide in the periphery. Mechanisms that drive heightened TCR sensitivity during development are not well characterized. Recently, it has been shown that the RAG complex has roles in addition to serving its important function in rearranging antigen receptor loci during VDJ recombination. RAG-mediated double stranded DNA breaks activate transcriptional programs that are essential for lymphocyte development. In this report I have utilized a mouse with a point mutation in the RING domain of RAG1 that inactivates RAG1's E3 ligase activity. I show that the RAG1 E3 ligase function is necessary for efficient positive and negative selection. Additionally, protein expression of the signaling molecules ZAP70, SYK, and CD80 is reduced in thymocytes in the absence of functional RAG1 E3 ligase activity. Disrupted T cell development due to absence of functional E3 ligase activity leads to hyperactive peripheral T cells that cause more severe immune-mediated pathology when transferred to $RAG1^{-/-}$ hosts in the absence of Tregs. I propose that RAG1 acts as

a biological clock to upregulate proteins involved in TCR signaling during T cell development in order to increase TCR sensitivity, enhance positive and negative selection, and protect from autoimmunity.

3.2: Introduction

The RAG complex is essential for the development of the vertebrate adaptive immune system through its role in rearranging antigen receptor loci encoding B and T cell receptors (BCRs/TCRs). Congruent with their importance, the RAG1 and RAG2 proteins that make up the RAG hetero-tetramer have been intently studied since their discovery ~30years ago (115, 130, 131). This process of VDJ recombination creates a diverse receptor repertoire that enables protection against a virtually infinite number of potential antigens, a necessary requirement for an adaptive immune response that is tasked with protection from a range of insults for which it has no prior knowledge (29, 37, 39). However, the outcome of the VDJ reaction is the formation of a random variety of receptors that vary from being non-functional, to self-reactive. For $\alpha\beta$ T cells, the thymus serves as a specialized organ in order to test functionality of and self-reactivity of developing T cells, and thus ensures that the mature, pre-immune T cell repertoire is both poised to offer immune protection without causing autoimmunity (58, 82-84, 132-134). Death of cells that cannot signal through their TCR is known as positive selection and ensures that TCRs are functional (67, 135). Similarly, thymocytes undergo negative selection whereby cells that signal too strongly on self-peptide MHC complexes encountered in the thymus are also deleted (67, 132). A conceptual difficulty lies in the problem that both positive and negative selection must occur by TCR:self-pMHC interactions. Since TCR signal strength is dictated by TCR:pMHC affinity(avidity) (136), this sets up a “goldilocks zone” of self-

reactivity between a non-functioning (weak) threshold, and an autoreactive, over-functioning (strong) threshold. The requirement for TCR selection on self-peptides results in a tenuous state where some over-functioning TCRs might not sufficiently signal deletion of cells harboring hazardous VDJ recombination products.

A number of mechanisms to overcome this tenuous state of potential failures in selection have been described, but most pertinent here is the suggestion that developing thymocytes are more sensitive to TCR stimulation than their mature counterparts (80-87). In 1991, Pircher et al. (86), showed that peptides of low affinity were able to induce clonal deletion of antigen-specific thymocytes, but the same antigen was not able to induce activation of peripheral T cells. This was expounded upon by the Hogquist group who showed nicely that thymocytes could signal more strongly on low-affinity antigen than peripheral T cells (80). Thus, this system of 'TCR tuning' sets up a scenario for more successful positive selection and more stringent negative selection, because self-peptides would be able to generate a stronger signal at both selection checkpoints (80-87). TCR detuning in the periphery then would limit autoreactivity because weak self-peptide signaling for positive selection is no longer needed, while strong self-peptide signaling (i.e., autoimmunity) would be attenuated by detuning (80-87). A number of regulators of this altered TCR sensitivity have been described, including Themis (137-139), Tespa1 (140), and miR181 (141). Despite the identification of the roles for these individual effectors, a global mechanism to drive thymic tuning is lacking.

For such a system to operate, it would be ideal for it to be under control of proteins present specifically during T cell development. Candidates such as the RAG1 and RAG2 proteins are perfectly poised for such a role because during $\alpha\beta$ T cell development, RAG

is expressed and active only in DN and DP thymocytes (63, 142, 143). Furthermore, it has been shown that RAG activity induces a transcriptional program that is essential for normal lymphocyte development, in addition to its role in rearranging lymphocyte receptors (92-94).

Beyond activation of a transcriptional response, the Bassing/Behrens labs have shown that the RAG1 N-terminal domain upregulates the post-transcriptional expression of a number of proteins (unpublished). The RAG1 N-terminus is known to bind a kinase that helps repair RAG DSBs, a ubiquitin ligase, and an RNA splicing factor (54-57); yet, potential functions of these interactions are unknown. This study primarily utilizes a mouse with a point mutation in the RING domain of RAG1 (P326G: $RAG1^{PG}$ mouse) that inactivates RAG1's E3 ligase activity (49-53). *In vitro* analysis of the P326G RAG1 mutant shows mildly defective VDJ recombination (53), but is shown to be necessary for ubiquitylation of histone H3 to enhance RAG cleavage of chromosomal DNA (49, 52). We utilize this mouse to show that the E3 ligase activity of RAG1 is required for efficient positive and negative selection of thymocytes and show that the expression of the signaling molecules ZAP70, SYK, and CD80, are negatively affected in the absence of RAG1 E3 ligase activity. We therefore propose a model whereby RAG1 activity, and specifically the N-terminal domain of RAG1, alters the thymic proteome in such a way to enhance thymocyte signaling. Thus, RAG1 acts as a biological clock in order to effectively tune TCR signaling in a way that is limited to developmental stages. Additionally, mutations in the N-terminal domain of RAG1 are associated with Omenn Syndrome, a disease which manifests as concurrent lymphopenia and autoimmunity (44, 96, 99), and

we show that peripheral T cells from RAG1^{PG} mice are hyper-functional, and can cause more severe and disseminated disease when transferred into RAG1$^{-/-}$ mice.

3.3: Results

RAG1^{PG} mice display T and B cell developmental defects at antigen receptor rearrangement steps

While the RAG1^{PG} mice are currently being investigated by two collaborative laboratories, and the baseline development of T and B cells in these mice had been described to us (Personal Communications: David Schatz – Yale; Craig Bassing – Children's Hospital of Philadelphia), since these mice have yet to be published on, I sought to investigate basic T and B cell development in these mice within our lab. In agreement with our collaborators, RAG1^{PG} mice have a mild reduction in double positive (DP) thymocytes (Figure 3-1 A/B), that is likely due primarily to a defect in the DN3 to DN4 transition (Figure 3-1 C/D). This is to be expected in the case of defective VDJ recombination, as observed in Chapter 2 in the RAG1^{NX} mouse, since this is when TCRβ is rearranged. Similarly, there is a defect in B cell development at the pro-B to pre-B transition where the analogous rearrangement of the Ig-heavy chain occurs (Figure 3-1 E-G). Interestingly, and in agreement with our collaborators, these developmental defects appear to influence total peripheral B cell numbers to a far higher extent than T cell counts (Figure 3-1 H). This chapter, however, will focus exclusively on T cells. Since developmental defects appear to occur at antigen receptor rearrangement stages, in collaboration with the Bassing Lab, I sought to determine whether VDJ recombination was impaired in developing thymocytes.

RAG1^{PG} thymocytes have reduced recombination efficiency

TCRβ rearrangements in RAG1^{PG} thymocytes were analyzed by TaqMan PCR as for the RAG1^{NX} mice (Figure 2-3) on sorted DN3 cells. There is a clear overall reduction in VDJ recombination in PG DN3 cells compared to WT counterparts (Figure 3-2 A/B). While less severe, TCRα rearrangement is also slightly reduced at some Vα-Jα pairings (e.g. Vα13-Jα17), but not at others (e.g. Vα14-Jα31) (Figure 3-2 C). Unlike analysis of Vβ-DJβ rearrangements, it is not possible to assay the entire spectrum of Vα-Jα pairs. Therefore, we cannot for certain say that overall recombination at the TCRα locus is reduced, especially considering all Vα14, Vα17, and Vα18 combinations assayed appear to be normal, and it is feasible that some pairings may be over-represented in the PG population. In fact, overall, surface expression of TCRβ at the DP pre-positive selection phase of T cell development (which requires expression of a correctly rearranged TCRα chain) is equivalent between WT and PG thymocytes (Figure 3-3 A). The most striking part of this analysis is the decrease in RAG1^{PG} cells that progress to Stage 2 of T cell development.

Positive and negative selection is impaired in RAG1^{PG} mice

In attempting to determine whether DP pre-positive selection thymocytes (CD8$^+$ CD4$^+$ CD69$^-$ CCR7$^-$) express equivalent TCRβ as a proxy for TCRα rearrangement efficiency (described above), I noticed a substantial decrease in cells in the first stage of positive selection (CD69$^+$ CCR7$^-$) (Figure 3-3 A). This suggested that PG thymocytes have a positive selection defect that is independent of TCRα rearrangement. As a proxy to establish whether positive selection is impaired in the PG T cells, I determined the

frequency of cells that fall within the positive selection/Stage 2 gate of total Stage 1 and Stage 2 cells that express surface TCRβ. Hence these cells *can* theoretically signal through their TCR and be positively selected (Figure 3-3 B). PG thymocytes appear to have a defect in their ability to positively select. It is important to note that this is an imperfect system for measuring positive selection as it cannot take into account differences in pace of progression from stage 2 to stage 3 and beyond. Nor deletion due to exceptionally strong TCR signals that may be occurring at this stage. More sensitive approaches including pulse-chase experiments should be utilized to better confirm positive selection defects. An improved system is discussed in more detail in Chapter 3.4. Despite this limitation, this data suggests a defect in positive selection of PG thymocytes. Given this, I was interested in determining whether there was also a defect in negative selection.

To assay whether $RAG1^{PG}$ thymocytes have a defect in negative selection, I utilized a naturally occurring system of super-antigen mediated negative selection (78, 144, 145). All laboratory mice express remnant MMTV retroviral products, and some of these retroviral products can bind the MHCII I-E^d protein and certain TCRvβ alleles, including vβ5.1-5.2 and vβ12, in a superantigen manner such that T cells expressing these alleles encounter an extremely strong TCR signal and are very efficiently deleted (78, 144, 145). BALB/c mice express the I-E^d molecule, while C57BL/6 (B6) mice (the background of $RAG1^{PG}$ mice), do not. Therefore vβ5.1-5.2 and vβ12 are efficiently deleted in BALB/c but not B6 mice. In order to test the effect of the PG mutation on deletion of 'autoreactive' TCRs, I crossed $RAG1^{WT}$ and $RAG1^{PG}$ mice to $RAG1^{-/-}$ BALB/c mice to generate F1 mice that were either $RAG1^{WT/-}$ I-$E^{d+/-}$ or $RAG1^{PG/-}$ I-$E^{d+/-}$. To examine whether there was a

difference in deletion of specific vβ alleles, I measured the expression of I-E^d:MMTV reactive (vβ5.1-5.2, vβ12) and I-E^d:MMTV unreactive (vβ6, vβ8, vβ14) TCRs on pre-selected (DP) or post-selection (SP, $CD24^{lo}$, $TCR\beta^{hi}$) thymocytes (Figure 3-3 C). "Escape" from negative selection was determined as follows: Of cells that create a TCR with a certain vβ allele (pre-selected cells), what is the frequency that progress through development to post-selection? This was normalized to relative frequencies of total pre- and post-selected cells (which as indicated in Figure 3-1 is slightly different between genotypes), and in order to combine data from independent experiments, this ratio was normalized to the WT mice in each experiment (Figure 3-3 D). It is clear that PG mice have increased 'escape' of vβ5.1-5.2 and vβ12 TCRs compared to WT, and in important controls, the development of nonreactive TCRs is not impacted. This is an incredibly interesting finding and the first evidence that RAG1 may have a role in controlling positive and negative selection. Because of these defects in selection, we were then interested in probing the mechanism as to why these defects exist.

SYK protein expression is decreased in DN3 thymocytes from $RAG1^{PG}$ mice

It is becoming increasingly clear that the N-terminal region of RAG1 has roles outside of its canonical role in VDJ recombination. One such function appears to be the control of protein expression, and in B cell systems, the N-terminal portion of RAG1 has been implicated in driving enhanced protein expression of the pro-survival genes Pim2 and Bcl-xl, and the important B cell signaling molecule SYK (Unpublished data, Bassing Lab, Personal Communication). While canonically associated with B cell receptor (BCR) signaling, SYK is also important in driving pre-T cell receptor signaling at the DN3 to DN4

transition (146). I therefore sought to determine whether SYK expression is impacted in RAG1^{PG} DN3 thymocytes via flow cytometry and show that there is an ~15% reduction in SYK protein expression compared to RAG1^{WT} DN3 cells (Figure 3-4 A/B). I then sorted DN3 cells to assess SYK mRNA expression and show that it is in fact even higher in RAG1^{PG} cells compared to WT counterparts (Figure 3-4 C). This is consistent with data from the Bassing Lab suggesting that in B cells, Pim2 and Syk expression is controlled by RAG1 at post-transcriptional stages (unpublished). Given the structural similarity between the signaling molecule SYK, and the canonical T cell signaling molecule ZAP70 (147), and the fact that TCR signal strength is intrinsically linked to positive and negative selection, I sought to determine whether ZAP70 was similarly affected in RAG1^{PG} thymocytes.

ZAP70 protein but not mRNA is decreased in developing RAG1^{PG} thymocytes and differences are restricted to the thymus

ZAP70 expression was assessed at all stages of positive and negative selection in thymocytes from WT and PG mice (Figure 3-5 A). ZAP70 expression increases following positive selection and a difference in expression of ZAP70 emerges after this stage with PG thymocytes having reduced expression compared to WT (Figure 3-5 B/D). This is independent of surface TCR expression (Figure 3-5 C/E). Like SYK, the effect of the PG mutation on ZAP70 protein expression is independent of mRNA levels, as ZAP70 mRNA in sorted thymocytes is not different between WT and PG cells (Figure 3-5 F).

Interestingly, the effect of the PG mutation on ZAP70 expression persists during thymocyte development, and after RAG1 expression has itself ceased. However, I sought to determine whether this difference was permanent by staining for ZAP70 in peripheral T

cells. Expression of ZAP70 is equivalent between WT and PG splenic T cells (Figure 3-6 A/B). Therefore, while the difference in ZAP70 persists throughout T cell development, this is not permanent. I next sought to probe further the mechanism by which the RAG1 E3 ligase domain may be impacting protein expression.

The N-terminal region of RAG1 controls protein expression by various mechanisms

This section attempts to answer two major mechanistic questions, and was performed in collaboration with Dr. Charline Miot from the Bassing Lab:

(1) Are RAG-mediated double stranded DNA breaks (DSBs) required for the difference in protein expression observed in PG thymocytes?

(2) What is the contribution of the E3 ligase activity of RAG1 in this control, compared to the rest of the N-terminal domain of RAG1.

In order to answer these questions, we developed a genetic system to interrogate the role of RAG1-DSBs, the N-terminal domain of RAG1, or the specific E3 ligase activity of RAG1. As detailed in Table 3-1, Artemis$^{-/-}$Bcl2^{tg} mice (mice that cannot repair RAG-mediated DSBs) were bred to express RAG1^{WT}, RAG1^{PG}(E3 ligase-dead), or RAG1^{core}(lacking entire N-terminal domain of RAG1). These were compared to RAG1$^{-/-}$(complete absence of RAG1, i.e. no DSBs nor RAG1 protein), and RAG1^{DA}(presence of RAG1 without DSBs) mice. Since none of these mice can rearrange antigen receptor loci, they all express the fully rearranged transgenic TCRβ, Vβ14^{NT} (referred to as VβNT)(148), which allows them to progress through the DN3→DN4 TCRβ-selection checkpoint, before they become trapped at the DP stage since they cannot successfully rearrange TCRα. Therefore,

thymocytes from all these cells do not fully rearrange a functional TCR, do not get a TCR-signal, and therefore maintain RAG expression. Thus, this system serves us well to answer the proposed questions.

We show that RAG1-enhanced ZAP70 protein expression does not require RAG-mediated DSBs since $RAG1^{DA}$ mice have equivalent ZAP70 protein expression to $RAG1^{WT}$ expressing thymocytes (Figure 3-7 A). However, the RAG1 N-terminal domain is essential for this activity since $RAG1^{-/-}$ and $RAG1^{core}$ expressing mice have substantially reduced ZAP70 expression (Figure 3-7 A). $RAG1^{PG}$ expressing mice in this experiment do not have an observable difference in ZAP70 expression, but this is expected since there is not an observable difference in ZAP70 expression in equivalent Stage 1 thymocytes between normal WT and PG mice (Figure 3-5 B/D). Hence, the RAG1-E3 ligase activity appears to contribute to but is not the only region of the N-terminal domain of RAG1 that enhances ZAP70 expression. As with the normal WT vs PG thymocytes, these differences in protein expression are not due to lower mRNA levels of ZAP70 (Figure 3-7 B).

We were interested to determine whether this mechanism of control was true for all proteins that RAG1 regulates expression of. We chose to also investigate whether CD80 expression was induced on thymocytes in a similar manner since CD80 has been shown to be upregulated following RAG-DSBs in B cells (92). Unlike ZAP70, CD80 expression on thymocytes is absolutely dependent on the presence of RAG-DSBs since $RAG1^{DA}$ expressing mice cannot upregulate CD80 expression (Figure 3-7 C). $RAG1^{PG}$ and $RAG1^{Core}$ mice also have decreased CD80 protein expression. This could either be due to reduced DSBs by these more inefficient RAG1 enzymes, or by direct regulation of protein.

Unlike ZAP70, CD80 protein expression directly correlates with CD80 mRNA expression (Figure 3-7 D). However, like the effects of the RAG1 N-terminal domain on ZAP70 expression, there is again a graded decrease from WT → PG → Core, suggesting that while the E3 ligase domain of RAG1 enhances ZAP70 and CD80 expression, other domains within the N-terminal region also contribute. Overall this data highlights that the expression of multiple proteins is enhanced by RAG1 activity, and the RAG1 N-terminal domain by a number of different mechanisms. It is therefore likely that the RAG1 protein can drastically reshape the proteome of a developing lymphocyte and have important effects on the outcome of lymphocyte development. This requires larger proteomic and transcriptomic analysis to delve further, but since the focus of this report is largely focused on the RAG1-mediated control of signaling proteins, I sought to investigate further whether TCR signaling in $RAG1^{PG}$ thymocytes was negatively impacted.

$OT1^{+}$: $TAP1^{-/-}$ bone chimeras do not provide an adequate system to investigate signaling differences between thymocytes

Given the small but consistent differences in ZAP70 expression within developing WT and PG thymocytes, we expect small differences in TCR signaling. While we attempted cross-linking of thymocyte TCRs with anti-CD3 to measure signal strength, this is likely too strong of a TCR signal and thus not sensitive enough compared to normal pMHC:TCR crosslinking that have avidities 3-6 orders of magnitude less than anti-CD3 crosslinking antibodies (149, 150). These experiments did not show consistent differences in signaling between WT and PG thymocytes. Therefore, we required the use of a fixed TCR to measure pMHC:TCR signaling of developing thymocytes and chose the well characterized ovalbumin-peptide (SIINFEKL) specific, MHCI restricted, OT1 TCR due to its vast array of

available tools. Unfortunately, in order to measure the effects of RAG expression, having a pre-rearranged TCR is not ideal as positive selection signals rapidly turn off RAG1 expression. As such, effects of RAG1 on ZAP70 expression are likely muted, and in fact this was the case as no difference in ZAP70 expression was observed between $OT1^{+}RAG1^{WT}$ and $OT1^{+}RAG1^{PG}$ thymocytes (data not shown).

One method to subvert these problems is to perform these experiments on a non-selecting background. We chose to begin with the $TAP1^{-/-}$ background as mice lacking TAP1 cannot load peptide onto MHCI, therefore surface expression of MHCI is very low and MHCI-restricted CD8+ T cells cannot be positively selected (151). Therefore, this should satisfy the requirements of the ideal system where OT1 expressing T cells are trapped at the DP Stage 1 phase of development and maintain high expression of RAG1.

In a first attempt to develop such a system, I chose to make mixed OT1 $RAG1^{WT}$ and OT1 $RAG1^{PG}$ bone marrow chimeras in $TAP1^{-/-}$ hosts (Figure 3-8 A). Unfortunately, the OT1 cells (both WT and PG) in $TAP1^{-/-}$ hosts are not effectively halted at DP-Stage 1 (CD69- CCR7-) (Figure 3-8 B/C). While the OT1 cells in $TAP1^{-/-}$ hosts are partially delayed compared to regular OT1 thymocytes (which progress through positive selection very rapidly), there is significant upregulation of CD69, indicating a positive selection signal that is likely sufficient to turn off RAG1 expression, and subsequently any RAG1 effects (Figure 3-8 B/C). This is striking when compared to $V\beta14^{NT}RAG1^{-/-}$ or $V\beta14^{NT}Artemis^{-/-}Bcl2^{tg}$ controls which are preventing from progressing though positive selection by virtue of their inability to rearrange $TCR\alpha$.

As a result, no difference in ZAP70 expression between OT1 RAG1^{WT} and OT1 RAG1^{PG} thymocytes is detected (Figure 3-8 D). Phospho-γH2AX, which is a unique histone that gets recruited to sites of DNA breaks, was used as a proxy for RAG-DSB break formation, and hence RAG1 expression (152). OT1 cells in TAP1$^{-/-}$ hosts show almost no increase in phospho-γH2AX staining compared to regular OT1 cells, suggesting that RAG1 expression and activity is not maintained in this system as expected (Figure 3-8 E). Vβ14^{NT}RAG1$^{-/-}$ and Vβ14^{NT}Artemis$^{-/-}$Bcl2^{tg} thymocytes were used as negative and positive controls for zero, and many RAG1-induced DSBs respectively (Figure 3-8 E, Table 3-1).

While ZAP70 is not observably different between WT and PG expressing OT1 cells, we were still curious to probe whether this system could be utilized to determine small TCR signaling differences by thymocytes. Thymocytes from chimeric mice were plated onto bone marrow derived dendritic cells (BMDCs) that had been pre-loaded with the canonical SIINFEKL peptide (for which the OT1 TCR is specific for) of different concentrations and altered peptide ligands (for which the OT1 TCR has various affinities for). CD69 expression was read out as a measure of TCR stimulation and despite there being no difference between OT1 RAG1^{WT} and OT1 RAG1^{PG} cells, this system should be sensitive enough to detect small signaling differences (Figure 3-8 F). This system is in the process of being improved and will be discussed further in Chapter 3-4.

ZAP70 overexpression cannot rescue positive and negative selection defects of RAG1^{PG} expressing T cells

The next question was whether overexpression of ZAP70 could rescue any of the developmental defects observed in the RAG1^{PG} mice. Namely:

(1) Can ZAP70 overexpression enhance the DN3 to DN4 transition for RAG1^{PG} thymocytes, since ZAP70 can compensate for SYK signaling in DN3 cells (146)?

(2) Can ZAP70 overexpression rescue the positive selection defect that RAG1^{PG} thymocytes display?

(3) Can ZAP70 overexpression rescue the negative selection defect that RAG1^{PG} thymocytes display?

To address this, hematopoietic progenitors (LSK cells) were sorted from WT and PG bone marrow, transduced with ZAP70 overexpressing (ZAP70^{OE}) retrovirus (or GFP control), and adoptively transferred into BALB/c X B6 F1 hosts (Figure 3-9 A). Since transduction of the donor cells was not 100%, and WT and PG cells were co-transferred into hosts, within a single host, 4 cell populations can be analyzed:

1. WT GFP^{-}
2. WT GFP^{+} (Control or ZAP70^{OE})
3. PG GFP^{-}
4. PG GFP^{+} (Control or ZAP70^{OE})

Successful engraftment of host mice with ZAP70^{OE} cells that can develop into thymocytes was achieved in a proportion of the mice, and ZAP70 expression is driven to a very high

level in these cells (Figure 3-9 B). In asking whether ZAP70 overexpression can enhance the DN3 to DN4 transition of developing PG T cells, I examined CD25 expression on $CD4^-$ $CD8^-$ $CD44^{lo}$ cells from GFP+ control and $ZAP70^{OE}$ thymocytes. There was a large amount of variation between samples, but the DN3→DN4 defect appears to be present in PG-control cells compared to WT-control cells (Figure 3-9 C). This trends towards being a less prominent defect in $ZAP70^{OE}$ cells (Figure 3-9 C). However, this needs to be repeated with large numbers of mice to account for the high variability.

To observe the effect of ZAP70 overexpression on positive selection, CD69, TCRβ, and CCR7, were measured on the surface of developing thymocytes (Figure 3-9 D). ZAP70 overexpression did not change the expression of any of these developmental markers. To account for differences in TCRβ expression between WT and PG cells, CD69 MFI was normalized to TCRβ MFI of Stage 1/2 (CCR7-), TCRβ+ thymocytes. The expectation was that $ZAP70^{OE}$ should increase the relative CD69 signal per amount of surface TCR, since TCR signaling should be increased. This was not the case and in fact the opposite occurred with the CD69/TCRβ ratio decreasing in $ZAP70^{OE}$ cells (Figure 3-9 E). Potentially, enhanced signaling may promote more rapid upregulation of CCR7 and therefore CD69 high cells quickly exit the Stage 1/2 gate, or such strong signaling may enhance negative selection at Stage 2. Of note, in this system there was not significant differences in CD69/TCRβ expression between WT and PG cells potentially suggesting that the positive selection defect may not be cell intrinsic.

Finally, we wished to determine whether negative selection was impacted by ZAP70 overexpression. In using BALB/c X B6 F1 hosts, we hoped to measure vβ5.1-5.2

and vβ12 deletion in these mice as described for Figure 3-3 (C/D). Unfortunately, in this chimeric system, neither vβ5.1-5.2 nor vβ12 (data not shown) were efficiently deleted (Figure 3-9 F). ZAP70 overexpression had no effect on deletion (Figure 3-9 F). Thus, it was not possible to effectively assess negative selection in this model. Overall this model was not convincingly able to answer any of the questions posed regarding whether ZAP70 overexpression can rescue any of the developmental defects observed by PG mice.

Peripheral RAG1^{PG} T cells can make more cytokines ex vivo and have a prolonged expansion of short-lived effector T cells following LCMV infection

Given the work by the Sun group showing the long-lived effects of RAG activity on the effector functions of NK and T cells *(153),* and that we expect the PG mice to have an increase in the number of peripheral T cells that are autoreactive, we were interested to probe their surface phenotype and cytokine producing ability *ex vivo*. Naïve and memory phenotypes of CD4+ and CD8+ T cells from naïve WT and PG mice are not exceptionally different (Figure 3-10 A/B). Upon stimulation with PMA and ionomycin, a slightly higher frequency of CD4+ and CD8+ T cells from naïve PG mice are able to make cytokines (Figure 3-10 C/D).

I thought it possible that these small differences may be amplified by an immune stimulus such as a viral infection. Therefore, WT and PG mice were infected with an acute strain of lymphocytic choriomeningitis virus (LCMV-Armstrong) and sacrificed at both an early timepoint at the peak of the immune response (Day 7 post-infection) and at a later, memory timepoint (Day 35 post-infection). At day 7 post-infection there is no difference in the numbers of CD4+ or CD8+ T cells between WT and PG mice, but by day 35 post-

infection a larger number of both CD4+ and CD8+ T cells remained, suggesting a defect in the contraction phase of the immune response (Figure 3-11 A). PG T cells were equally able to mount an antigen specific CD8+ T cell response as measured by gp33-tetramer staining and peptide-specific re-stimulation (data not shown). Contraction of antigen-specific T cells also appeared to be equivalent between WT and PG genotypes (Figure 3-11 B, data not shown).

The most striking difference between the WT and PG CD8+ T cells was the enhanced retention of the short-lived effector (SLEC) population at day 35 post-infection (Figure 3-11 C). While important for clearance of infection early, these cells typically contract substantially after the infection has been cleared (as observed by the WT cells). Correspondingly, PMA/ionomycin re-stimulation of cells isolated from these mice shows that an increased frequency of PG CD8+ CD44+ T cells make IFNγ and TNFα at day 35 but not day 7 post-infection (Figure 3-11 D). While this is not a large difference in cytokine production, this suggests that activated effector PG T cells do not contract as efficiently as WT counterparts.

Naïve PG CD4+ T cells induce more severe colitis and disseminated inflammation when transferred into RAG1$^{-/-}$ hosts

While the evidence suggesting that negative selection of RAG1^{PG} T cells is impaired is strong, in such a setting it would be expected that escape of autoreactive T cells should lead to autoimmunity. However, as discussed in Chapter 2, C57BL/6 mice are exceptionally resistant to autoimmune diseases such as AIRE deficiency (126). Similar to the RAG1^{NX} mice, RAG1^{PG} mice do not present with overt autoimmune symptoms (data

not shown). In an attempt to determine whether $RAG1^{PG}$ T cells can cause autoimmunity, I utilized the model of autoimmune colitis which involves the transfer of naïve, non-Treg CD4+ T cells into $RAG1^{-/-}$ hosts (154, 155). This system, because of the absence of Tregs, leads to the induction of CD4+ T cell driven colitis (156). The hypothesis for this experiment is that if $RAG1^{PG}$ mice have more autoreactive CD4+ T cells than $RAG1^{WT}$ mice, then transfer of PG cells to $RAG1^{-/-}$ hosts should cause more severe colitis and/or disseminated autoimmunity.

The first observation was that there was a wide spread of phenotypes from host mice that got either WT or PG cells. Despite this a number of differences emerged. PG-transfer mice had slightly accelerated weight-loss compared to WT-transfer mice (Figure 3-12 A). Both WT and PG mice lost weight compared to co-housed “No transfer”(NT) controls, indicating both WT and PG CD4+ T cells can cause disease in this model, in our animal facility (an important consideration given the effect that microbiota and animal facility can have on the outcomes of this model (157, 158)). Death in this model within 8 weeks post-transfer of cells is unexpected, and even though it did not reach statistical significance, more PG-transfer mice died prematurely (Figure 3-12 B). Additionally, there was clear signs of non-colitis pathology that was present primarily in the PG-transfer mice, including three mice with very severe dermatitis (Table 3-2). Overall there were 10 (of 18 total) adverse, non-colitis pathologic events observed for the PG-transfer mice compared to 1 (of 15 total) WT-transfer mice.

Gross characterization of colitis severity included measurement of colon shortening, where PG-transfer mice fare worse than WT-transfer mice (Figure 3-12 C); and colon wall thickening (measured as the ratio between weight and length of colon) is

equivalent between WT- and PG-transfer mice, but worse than NT controls (Figure 3-12 D). At the time of writing this report, pathology photographs are unavailable, yet upon blinded examination of hematoxylin and eosin (H&E) stained colons from WT and PG mice, colonic tissue damage was worse in the PG-transfer mice. Pathology of colons, skin, livers, and lungs is forthcoming.

Strikingly, spleen weights of the PG-transfer mice were substantially larger than the WT-transfer mice (Figure 3-12 E). This correlated with higher CD4+ T cell counts in the spleen of PG-transfer mice compared to WT-transfer, a similar trend in livers, but no difference in the mesenteric lymph nodes (MLN) (Figure 3-12 F). This suggests that WT and PG cells can equivalently expand in the gut and gut draining lymph node (i.e. MLN), but PG cells have an advantage in peripheral dissemination. Interesting to note, 0.5million transferred CD4+ T cells have the ability to expand in a lymphopenic host to a striking degree (up to 40million cells in one spleen).

The transfer model of colitis is driven in large part by Th1 and Th17 skewed cells (159, 160), hence I was interested to examine whether there were differences between WT and PG cells in their ability to make these cytokines and skew appropriately. There is a wide distribution of response mouse-to-mouse, cage-to-cage, and experiment-to-experiment. For example, in one experiment, there were vastly different abilities between WT and PG CD4+ T cells to make cytokines (Figure 3-12 G), where splenic and MLN isolated CD4+ T cells from PG-transfer mice made substantially more IFNγ, and less IL-17A upon *ex vivo* re-stimulation (Figure 3-12 G; data not shown). Despite this experiment to experiment variability, using mixed-effects methods to account for experiment as a

random variable, there is a statistically significant increase in IFNγ, and decrease in IL-17A production by spleen-isolated PG-transfer cells compared to WT (Figure 3-12 H).

Next the expression of the canonical Th1 and Th17 lineage defining transcription factors, Tbet and RORγt respectively, within the WT- and PG-transfer CD4+ T cells was examined. Like cytokine production, these results were variable with large experiment to experiment variation. From one experiment, Tbet expression from spleen and MLN isolated PG-transfer cells was far higher than in WT counterparts (Figure 3-12 I), yet when combined from multiple experiments there appears to be no overall difference in Tbet expression between WT and PG cells (Figure 3-12 J). For RORγt however, within and across experiments, a decreased frequency of PG-transfer CD4+ T cells express RORγt compared to WT counterparts (Figure 3-12 L). Importantly, there was no difference in induced Tregs between WT- and PG-transfer cells (less than 1%, data not shown). Normally, a large discussion of variability between mice and experiments would not be warranted, yet because of the mechanism of how we see inefficient negative selection within the RAG1^{PG} mouse occurring, such variation is not surprising. This will be expanded upon in Chapter 3.4.

Finally, since a higher frequency, and therefore a higher number of PG-transfer cells were able to make IFNγ following re-stimulation, I queried whether this led to a systemic increase in IFNγ. Serum ELISAs from these mice confirmed that serum IFNγ was increased in PG-transfer compared to WT-transfer mice (Figure 3-12 M). There was no statistical difference in serum IL-6 (Figure 3-12 N) between WT- and PG- transfer mice although there were two PG-transfer mice that had extremely high IL-6 levels

comparatively. As expected, the NT control mice had limited detected serum IFNγ or IL-6. IL-17A serum levels were not available at the time of writing but are forthcoming.

PG CD4+ T cells can skew to different effector fates equivalently in vitro, and outcompete WT cells, likely via enhanced survival

The difference between the effector fates of WT- and PG-transfer cells in the RAG1$^{-/-}$ transfer model was interesting. I hypothesized that that mechanism for this could be one of three, non-mutually exclusive reasons:

1. PG CD4+ T cells have an intrinsic epigenetic programming that favors Th1 skewing over Th17 skewing.
2. PG CD4+ Th1 cells outgrow WT CD4+ Th1 cells but not Th17 (hence the increase in total PG CD4+ T cells within the RAG1$^{-/-}$ host).
3. A higher frequency of PG CD4+ T cells can access tissue sites within the RAG1$^{-/-}$ host that permit environmental-dependent Th1 skewing over Th17 skewing.

To address options 1) and 2), naïve WT and PG CD4+ T cells were sorted and co-cultured with irradiated feeder splenocytes with anti-CD3 and anti-CD28 antibodies, along with relevant skewing cytokines for Th0, Th1, Th2, Th17, and Treg differentiation for 5 days. CD4+ T cells were then stained for relevant Th-defining transcription factors and re-stimulated with PMA/ionomycin to measure expression of Th-defining cytokines.

As displayed in Figure 3-13 (A-D), the large majority of the cells can make TNFα suggesting all cells at the end of the culture are functional. Importantly, Th1 cells make IFNγ (Figure 3-13 B), and Th17 cells make IL-17A (Figure 3-13 D). Overall, PG

CD4+ T cells appear to be slightly better than WT CD4+ T cells at making some of the cytokines in these conditions, and there is not a cell-intrinsic inability to skew towards Th17, which was the major question for this experiment. One of the more interesting and unexplained outcomes is the increased expression of TNFα by PG skewed Tregs compared to co-cultured WT Tregs (Figure 3-13 E). This was not due to an inability to become a Treg as PG Tregs express FoxP3 equivalently to WT Tregs. This is an interesting point for future investigation.

Strikingly, in support of option 2) above, PG cells were able to extensively outcompete WT cells in Th0, Th1, and Th2 conditions, but not Th17 or Treg conditions (Figure 3-13 G). To address whether PG CD4+ T cells could proliferate better than their co-cultured WT counterparts, a similar experiment was performed where sorted WT and PG cells were stained with CFSE and stimulated with anti-CD3/CD28 dynabeads to measure proliferation ability. There was no difference in proliferation between co-cultured WT and PG CD4+ T cells, suggesting that the major difference driving PG cells outcompeting WT cells *in vitro* is due to enhanced survival of PG cells. Whether this is the case *in vivo* is a point of future investigation, but increased survival of Th1 PG cells but not Th17 PG cells compared to WT counterparts could account for the difference seen in the RAG1$^{-/-}$ transfer model. This does not exclude the possibility that option 3) is also a contributing factor, which fits into the proposed model of RAG1's impact on negative selection and will be discussed below.

3.4: Discussion

In this report I describe for the first time in detail, the effect of a specific point mutation in RAG1 that eliminates RAG1 E3 ligase activity on *in vivo* T cell development. While the PG mutation in RAG1 has a modest decrease in VDJ recombination *in vitro* when RAG1 protein is in limiting quantities (53), this E3 ligase activity appears to be more important for the *in vivo* recombination activity of RAG1. This is likely due to the role that this E3 ligase domain has in ubiquitylating histone H3, and making RAG1 more permissible to chromatin, a role that is of greater importance when recombining chromosomal DNA compared to that used in the original recombination assays (49, 53). This defect in recombination causes a defect in TCRβ rearrangement at the DN3 stage of development, and likely a mild defect at TCRα as well (although this cannot be resolved completely with our data and would likely require single-cell TCR sequencing of unselected DP thymocytes to confirm). Despite this, the effect on total numbers of thymocytes is modest, and does not impact peripheral T cell counts, or phenotype. Peripheral T cell phenotype is important to consider, as an increase in memory phenotype cells would suggest T cell lymphopenia and homeostatic expansion as with the RAG1^{NX} mice. However, this is not the case suggesting lymphopenia driven expansion does not account for the sufficient numbers of T cells in the periphery. Strikingly, there is a far more severe effect on B cells. Perhaps this is a feature of pre-B cells having a relatively small proliferative burst compared to T cells from the DN → DP stage, hence B cells are less able to overcome small disturbances in recombination efficiency (68, 161, 162). Why this is the case for the RAG1^{PG} mice, but not RAG1^{NX} or RAG1^{core} mice is unclear. More in depth analysis of B cell development in the RAG1^{PG} mouse is being undertaken by our collaborator, David Schatz (Yale).

The most striking finding in this report is the effect that a point mutation in RAG1 has on positive and negative selection of T cells. While many mutations in RAG1 (and other genes that are important for T and B cell development) have been associated with Omenn's Disease (44), it is not clear how these mutations lead to autoimmunity. For the first time I present a direct effect of a RAG1 mutation on positive and negative selection. While complete mechanism has not been elucidated, it is intriguing to hypothesize that the N-terminal region of RAG1 is able to increase the expression of signaling molecules, such as ZAP70 and CD80, specifically during T cell development, such that TCR signaling of thymocytes is enhanced against specific peptides, and thymocytes that express auto-reactive TCRs are efficiently deleted. This increase in TCR signaling sensitivity in the thymus, as has been proposed by others (86, 163, 164), allows for efficient deletion of autoreactive cells, and cells close to the autoreactive threshold (Figure 3-14 A). Thymocytes that are positively selected on self-peptides in the thymus, will no longer be able to respond to these same peptides in the periphery (Figure 3-14 A). In the absence of such a delta in TCR signaling between thymocytes and peripheral T cells, more autoreactive cells can escape (Figure 3-14 B). Thus, this report proposes that since RAG1 expression is tightly regulated to T cells that are in the thymus (142, 165-167), and just prior to positive and negative selection, RAG1 acts as a biological clock to shape the thymic proteome to enhance thymic TCR signaling, only in the thymus. Importantly, the difference in ZAP70 expression is maintained until the stage of development where thymocytes have upregulated CCR7 and migrated to the medulla where most negative selection occurs (i.e. Stage 4)(67). It would be interesting to see how long these effects of RAG1 persist, and whether recent thymic emigrants (RTEs) maintain these effects given these cells are also hypersensitive to low affinity antigen (168). This experiment could

easily be performed using the RAG2-GFP reporter mouse that marks RTEs via continued expression of GFP for a period of time after RAG2 expression has ceased (168, 169).

One potential criticism of these data is that these are small differences in ZAP70 expression, and whether these differences are biologically significant is not certain. This is indeed fair criticism given that I was unable to correct the defects in positive and negative selection by $RAG1^{PG}$ T cells with overexpression of ZAP70. While it is very likely that RAG1 controls the expression of a large number of proteins, and ZAP70 alone may not be sufficient to rescue defects, other groups have shown that suboptimal TCR signaling due to hypomorphic ZAP70 leads to similar defects in negative selection and pre-disposition to autoimmunity (77, 79, 170). We also show a defect in negative selection in $RAG1^{PG}$ mice, on the order of magnitude that is similar to other reports.

This report also does not address whether this decrease in ZAP70, or indeed whether TCR signaling of thymocytes regardless of ZAP70 expression levels, is different in $RAG1^{WT}$ vs. $RAG1^{PG}$ cells as is proposed. Unfortunately, the system utilized here was imperfect in a number of ways, including that positive selection in $TAP1^{-/-}$ mice is slightly leaky such that there is not a complete absence of CD8+ T cells in these mice (171). Additionally, since this experimental system utilized chimeric mice, the OT1 $RAG1^{WT}$ and OT1 $RAG1^{PG}$ donor cells have expression of MHCI and may be able to induce positive selection of neighboring thymocytes. Indeed, we have since shown that thymocytes are able to present peptide and activate other thymocytes (data not shown). It is also possible that OT1 TCRs can positively select on MHCII, and in these mice it is possible to detect OT1 CD4 single positive thymocytes. In order to improve this model, OT1 $RAG1^{WT}$ and OT1 $RAG1^{PG}$ mice have been crossed to $\beta 2m^{-/-}$ mice, which should provide a more

complete block of positive selection, and continued RAG1 expression, and has been utilized previously to detect small changes in TCR signaling, positive selection in response to low affinity peptides, and negative selection in response to high affinity peptides (172).

While intriguing ideas, the presented data does not address mechanistically how the RAG1 N-terminal domain impacts expression of certain proteins, and many future experiments will be performed to address these problems. Additionally, it is unlikely that ZAP70, SYK, and CD80 are the only proteins affected by RAG1 activity. However, given that RAG-driven DSBs induce transcriptional upregulation of proteins such as CD80 which is further amplified by the RAG1 N-terminal domain, and that the RAG1 N-terminal domain is required only for the post-transcriptional expression of other proteins such as SYK and ZAP70, it appears the N-terminal domain of RAG1 shapes the proteome of the developing thymocyte in a number of ways (Figure 3-15). Regarding the post-transcriptional regulation of proteins such as SYK and ZAP70, it is tempting to postulate that the E3 ligase activity acts to stabilize such proteins via direct ubiquitylation. However, given that RAG1 is strongly localized to the nucleus, it is unlikely to have much contact with such cytoplasmic proteins. Therefore, either an intermediate effector that RAG1 can interact with performs this function, or the N-terminal domain of RAG1 is required for transcriptional regulation of a program of proteins that can allow for increased expression of other proteins, via enhanced translation, decreased degradation, or both. Experiments to address whether protein translation is increased, protein degradation is reduced, and proteome/transcriptome-wide analysis to measure the global effect of the RAG1 N-terminal domain, and specifically the E3-ligase activity, are planned for the immediate future.

Given we have shown a defect in negative selection by RAG1^{PG} T cells in a model of superantigen-mediated negative selection, it would be expected that mice may develop signs of autoimmunity. RAG1^{PG} mice do not present with overt autoimmunity, but this is not necessarily unexpected as C57BL/6 mice are relatively resistant to autoimmunity compared to other mouse strains, as discussed earlier. To address whether the expression of RAG1^{PG} causes a pre-disposition to autoimmunity, we have engineered the P326G mutation onto the non-obese diabetic (NOD) mouse background. This background presents with spontaneous autoimmunity due to impaired peripheral tolerance mechanisms (173-175). We expect to see RAG1^{PG} NOD mice present with disease more rapidly given the defect in negative selection/central tolerance. These experiments are forthcoming. Nevertheless, we were able to show that peripheral T cells had heightened cytokine producing capabilities at baseline, and that upon infection with LCMV, there was a persistent expansion of activated short-lived effector cells. In NK cells and CD8+ TCR transgenic T cells, it was described that RAG-DSBs were important in allowing these cells to become effective memory cells that persist after clearance of infection (153). This is similar to the phenotype presented here, as NK cells without RAG-DSBs were hyper-functional at baseline also. It was presumed that these lasting effects were a result of RAG-mediated DSBs that enhance survival later in the lifetime of the NK cell, yet it could be possible that the RAG1 N-terminal domain could act to provide these long-term benefits in the absence of DSBs. While RAG1^{PG} T cells appear to have enhanced survival, at least in terms of numbers at later time-points post LCMV infection, these models are slightly different, because RAG1^{PG} cells have still experienced RAG-DSBs (since they have a recombined TCR), and it is possible that survival, and hyper-functionality are mutually

exclusive. Future investigation will determine the role of RAG-DSBs vs. the N-terminal domain of RAG1 on these phenotypic effects.

Finally, in an attempt to determine whether $RAG1^{PG}$ T cells are able to cause autoimmunity, we utilized the model of adoptive T cell transfer into $RAG1^{-/-}$ hosts. This model removes the peripheral tolerance afforded by Tregs and leads to T cell driven colitis. For WT-transfer mice, this is typically limited to intestinal inflammation, most likely due to the inflammatory environment, and the abundance of non-self-antigen. While most mice develop some level of disease/inflammation, measured by weight-loss (compared to no transfer controls), colon shortening/thickening, or elevated serum IL-6 levels, the most obvious difference between $RAG1^{WT}$ and $RAG1^{PG}$ mice was the appearance of non-colitis associated disease phenotypes. These phenotypes were diverse, and this is to be expected when considering the model. It is hypothesized that $RAG1^{PG}$ mice have a defect in negative selection, and hence central tolerance, thus $RAG1^{PG}$ mice have more autoreactive T cells in the periphery. Yet the specificity of those T cells is likely to be extremely variable. Given 0.5million naïve CD4+ T cells are transferred into each of the $RAG1^{-/-}$ hosts, and there is on the order of 0.01-0.001% of cells that are specific for a given peptide (176), it is unlikely that all recipients will be transferred with autoreactive T cells that are reactive to the same tissue. Thus, it is maybe not surprising that the phenotypes observed are varied. Likewise, tissue distribution is important in specifying the effector differentiation phenotype of CD4+ T cells as different Th subsets are differentiated based on exposure to different cytokines (177), and different tissues have different expression of different Th skewing cytokines. Correspondingly, there is an altered number of $RAG1^{PG}$ cells than can make the Th1-defining cytokine, IFNγ, vs the Th17-defining cytokine, IL-

17A, potentially hinting at the possibility that $RAG1^{PG}$ T cells were able to access tissue sites that provided enhanced Th1 skewing conditions. While I showed that $RAG1^{PG}$ T cells do not have an intrinsic defect in becoming Th17 cells, $RAG1^{PG}$ Th1 cells outcompete $RAG1^{WT}$ Th1 cells, but the same is not true for the same cells skewed towards Th17. This enhanced expansion was not due to increased proliferation and is likely due to enhanced survival in Th1 conditions, however, further work is necessary to prove this. This phenotype also helps to explain why $RAG1^{PG}$ mice show equivalent early expansion of CD4+ and CD8+ T cells during LCMV infection (a strong Th1-skewing infection), but increased numbers of cells later after infection has presumably cleared.

Overall, this work provides an important step in the understanding of how RAG1 can modulate T cell development in addition to its well characterized role in VDJ recombination. This report opens up a wide array of possible future directions and areas of study, many of which are already underway.

3.5: Figures

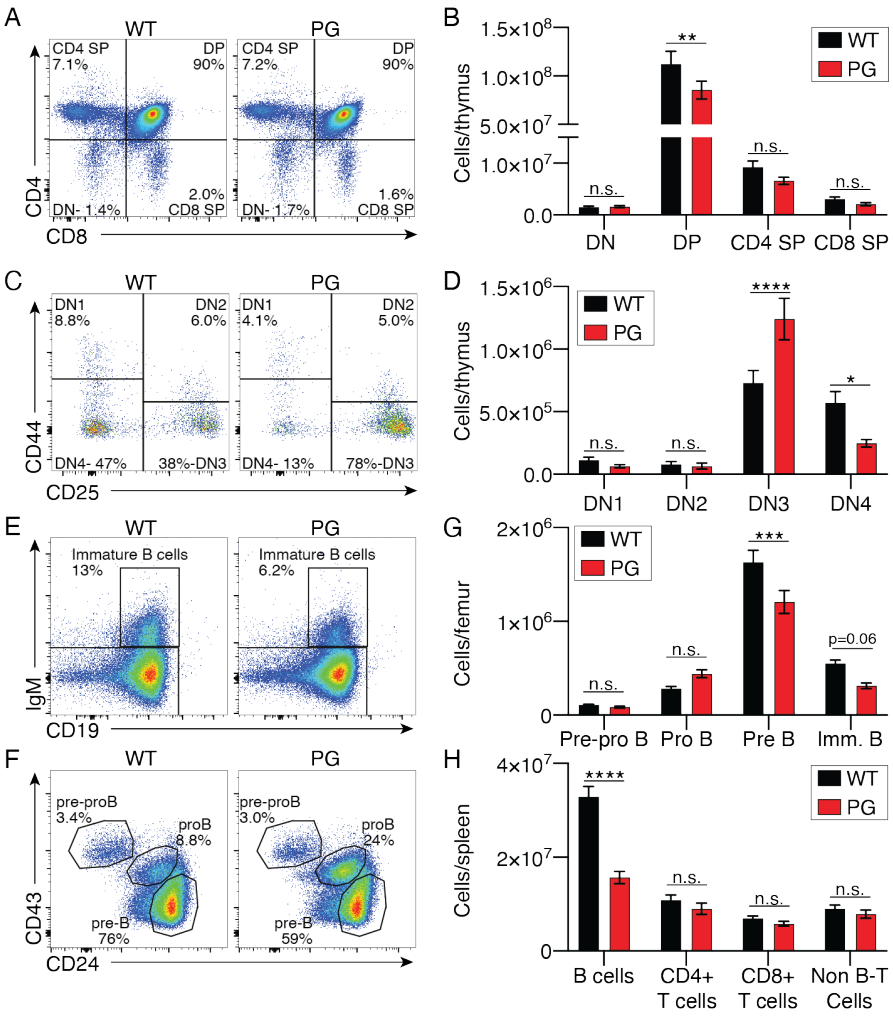


Figure 3-1: T and B cell development is impaired at antigen receptor rearrangement steps in RAG1^{PG} mice

Thymi from 6-9 week old RAG1^{WT} and RAG1^{PG} mice were analyzed for gross defects in T cell development. Frequencies **(A)** and counts **(B)** of DN, DP, and CD4 SP or CD8 SP cells shown. Pre-gated on live singlets. Frequencies **(C)** and counts **(D)** of relative DN populations (DN1-4) are shown. Pre-gated on live singlets, CD4$^-$, CD8$^-$, TCRβ^{lo}. Bone marrow from femurs was stained to interrogate B cell development. **(E)** Representative frequencies of IgM$^+$CD19$^+$ immature B cells of B220^{int}CD93$^+$ developing B cells. **(F)** Representative frequencies of pre-pro B, pro B, and pre B cells of IgM$^-$ developing B cells. **(G)** Quantification of total numbers of developing B cells from a single femur. 8-10 mice per genotype. **(H)** Number of splenic B (CD19$^+$TCRβ^-), CD4$^+$ T cells (CD19$^-$ TCRβ^+CD4$^+$), CD8$^+$ T cells (CD19$^-$TCRβ^+CD8$^+$), and non T or B cells (CD19$^-$TCRβ^-). 6 mice per genotype. All data combined from at least 2 independent experiments. Bars indicate mean +/- SEM. Stats: 2-way ANOVA with Sidak multiple comparison test.

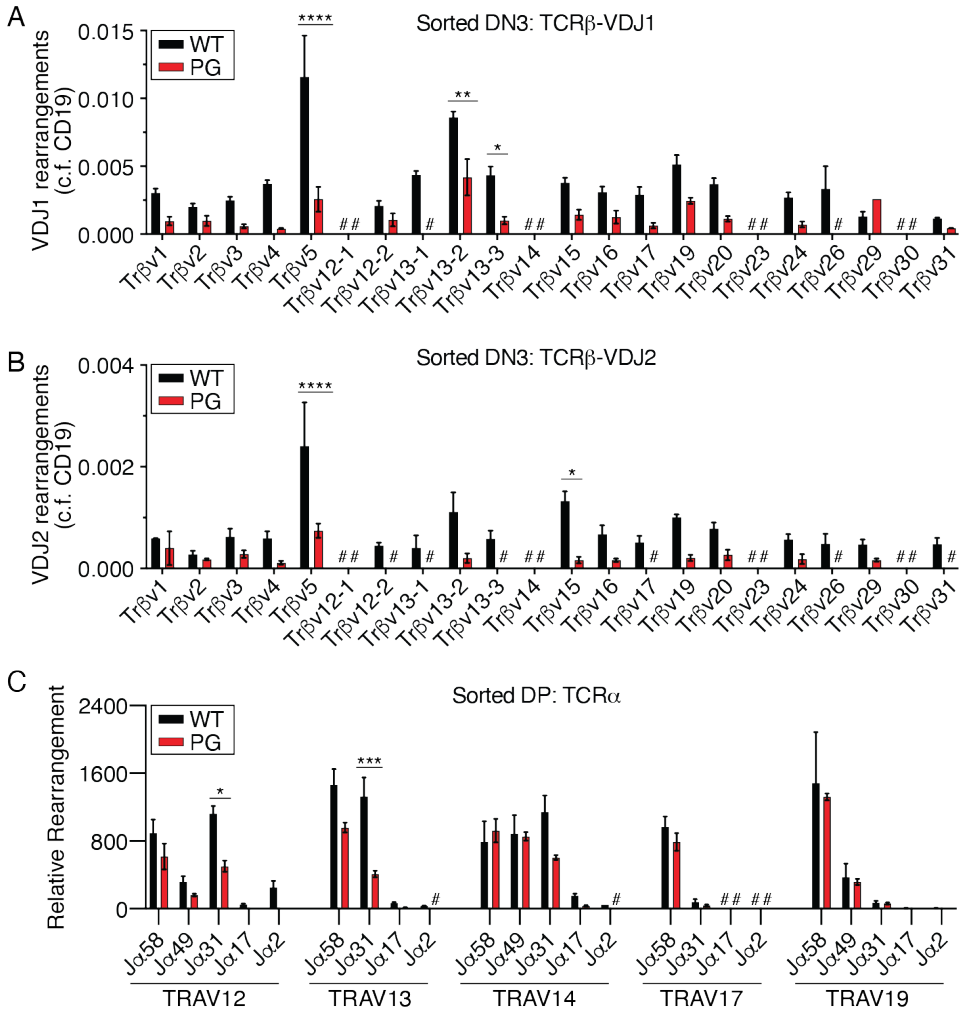


Figure 3-2: RAG1^{PG} thymocytes have reduced recombination efficiency

Sorted DN3 thymocytes from WT and PG mice were assayed by qPCR for frequencies of respective Vβ to DβJβ1 rearrangements **(A)**, and Vβ to DβJβ1 rearrangements **(B)**, 3 mice per genotype. $^{\#}$not detected. Gating: live, singlets, dump^{-} (B220, CD8, CD4, CD11b, CD11c, Gr1, NK1.1, Ter119), CD4^{-}, CD8^{-}, CD25^{+}, CD44^{-}. **(C)** TCR Vα to Jα rearrangements from sorted double positive thymocytes. Showing respective Vα to Jα pairings. Gating: Live singlets, CD4^{+}, CD8^{+}. 3 mice per genotype. $^{\#}$not detected.

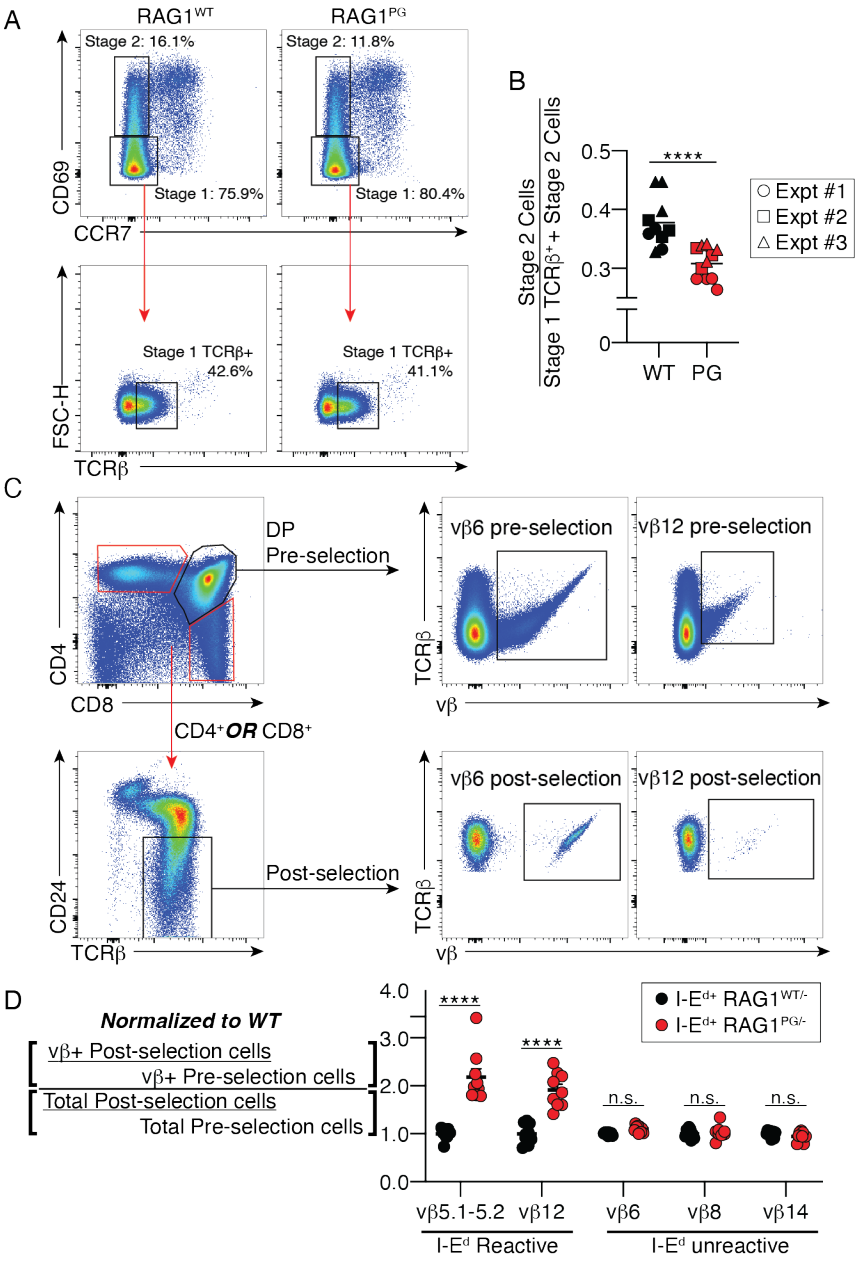


Figure 3-3: Positive and negative selection is impaired in RAG1^{PG} mice

(A) Gating strategy for measuring positive selection of thymocytes. Pre-gating: Live singlets, CD4$^+$ *AND/OR* CD8$^+$. **(B)** Number of stage 2 (CD69$^+$CCR7$^-$) cells of total Stage 2 and Stage 1, TCRβ$^+$ cells. 10-11 mice per genotype. Data combined from 3-independent experiments. Cross-bar = global mean. Stats: 2-way anova, indicating effect of genotype. **(C)** Example gating strategy of thymocytes from F1 generation of BALB/c RAG1$^{-/-}$ mice crossed with B6-RAG1$^{WT/WT}$ or B6-RAG1$^{PG/PG}$ mice, stained for vβ5.1-5.2, vβ6, vβ8, vβ12, and vβ14. Pre-gated on live singlets. **(D)** Ratio of vβ cells that progress to post-selection, of those cells that are produced at the pre-selection stage. Normalized to total thymocytes in either (pre or post-selection) population and WT controls.

Data combined from 2-independent experiments. 9 mice per genotype. Bars indicate mean +/- SEM. Stats: 2-way ANOVA with Sidak multiple comparison post-test.

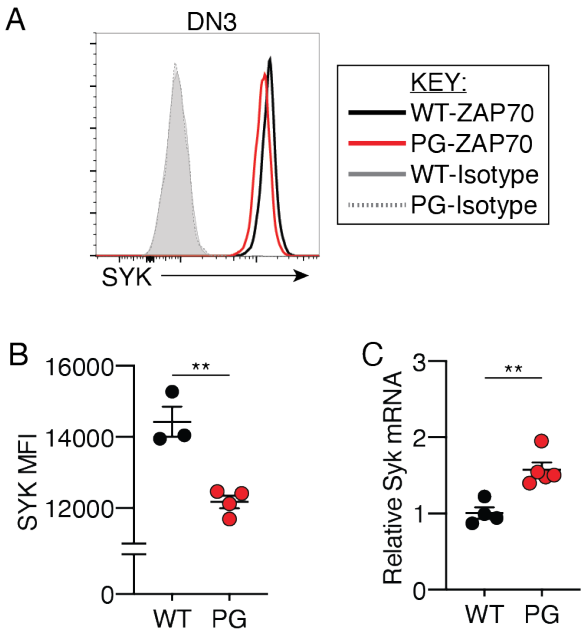


Figure 3-4: SYK protein, but not RNA expression is decreased in RAG1^{PG} DN3 thymocytes compared to RAG1^{WT}

(A-B) DN3 thymocytes were stained for total SYK protein expression. 3-4 mice per group. Pre-gating: Live singlets, CD4^{-}, CD8^{-}, TCRβ^{lo}, CD25^{+}, CD44^{-} (example gating seen in Figure 3-1). Representative of 2-independent experiments. **(C)** DN3 thymocytes were sorted and RNA isolated. RNA was converted to cDNA and Syk RNA expression was quantified. Relative RNA was calculated using ddCt method with HPRT used as housekeeping control, and a WT sample as control. Bars indicate mean +/- SEM. Stats = unpaired t-test.

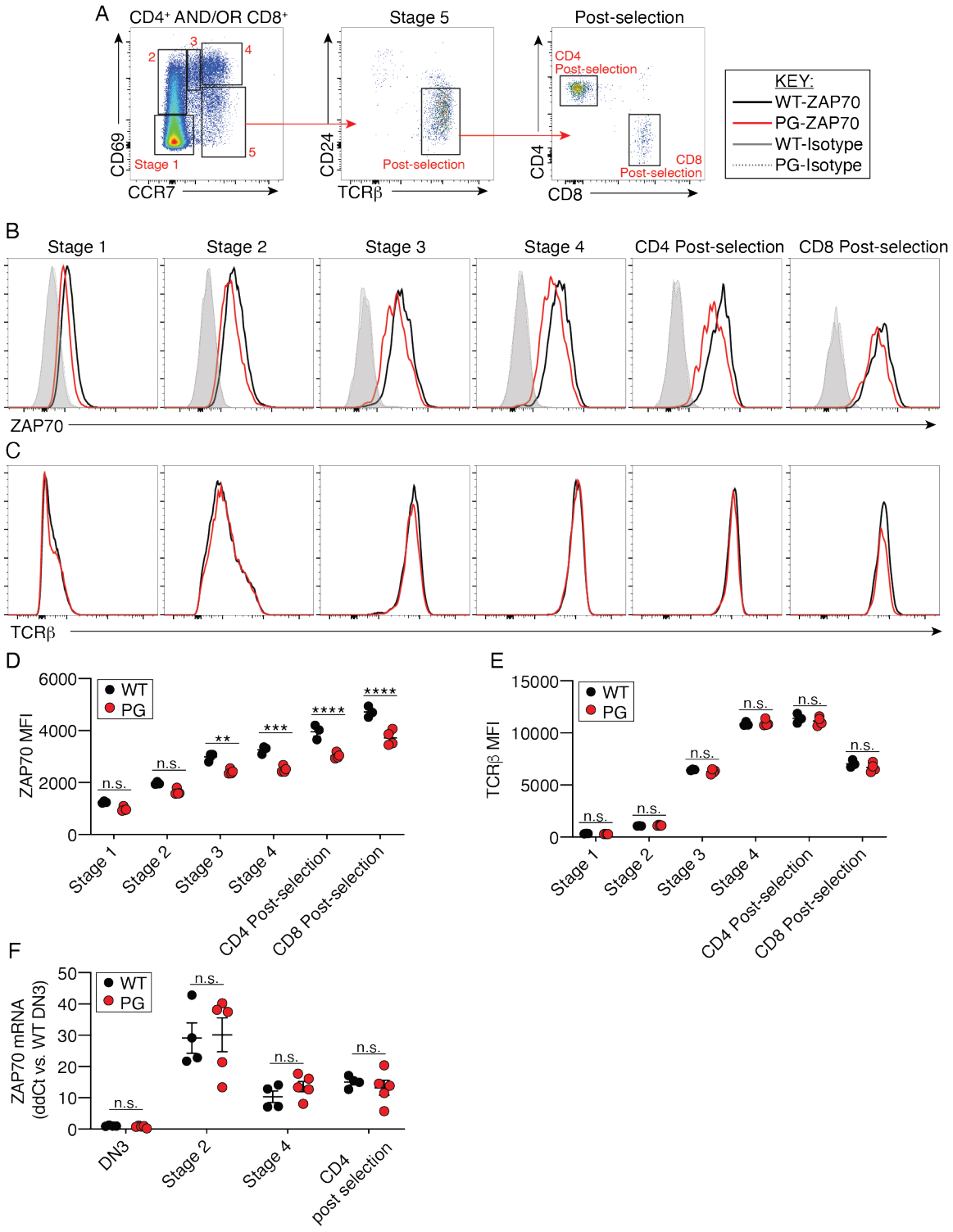


Figure 3-5: ZAP70 protein, but not mRNA, expression is decreased in developing RAG1^{PG} thymocytes

(A) Gating strategy for positive and negative selection stages. Pre-gating: Live singlets, $CD8^+$ *AND/OR* $CD4^+$. **(B)** ZAP70 staining on respective developmental stages. **(C)** TCRβ expression on respective developmental stages. **(D)** ZAP70 median fluorescence intensity on respective developmental stages. 4-5 mice per genotype. Representative of 3 independent experiments. **(E)** TCRb median fluorescence intensity on respective T cell developmental stages. 4-5 mice per genotype. Representative of 3 independent experiments. **(F)** Relative ZAP70 mRNA expression from respective sorted thymocyte populations. Relative RNA was calculated using ddCt method with HPRT used as housekeeping control, and a WT DN3 sample as control. Bars indicate mean +/- SEM. Stats = 2-way ANOVA with sidak multiple comparison post-test.

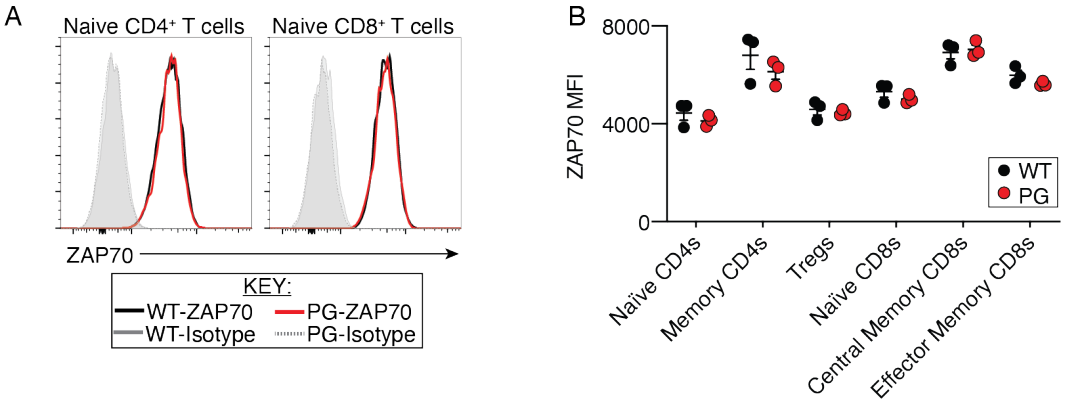


Figure 3-6: ZAP70 expression is not different in peripheral RAG1^{PG} T cells

(A) Representative histograms showing expression of total ZAP70 between RAG1^{WT} and RAG1^{PG} naïve CD4$^+$ T cells (left) and naïve CD8$^+$ T cells (right) compared to isotype controls. **(B)** Summary data of ZAP70 expression in Naïve CD4$^+$ T cells (CD44$^-$ CD62L$^+$), Memory CD4$^+$ T cells (CD44$^+$ CD62L$^-$), Tregs (CD25$^+$), Naïve CD8$^+$ T cells (CD44$^-$CD62L$^+$), Central Memory CD8$^+$ T cells (CD44$^+$CD62L$^+$), and Effector Memory CD8$^+$ (CD44$^+$CD62L$^-$). Pre-gating: Live singlets, TCRβ^+, CD4$^+$ *OR* CD8$^+$. 3 mice per genotype, representative of N=2 independent experiments. Stats: 2 way ANOVA with Sidak multiple comparison post-test.

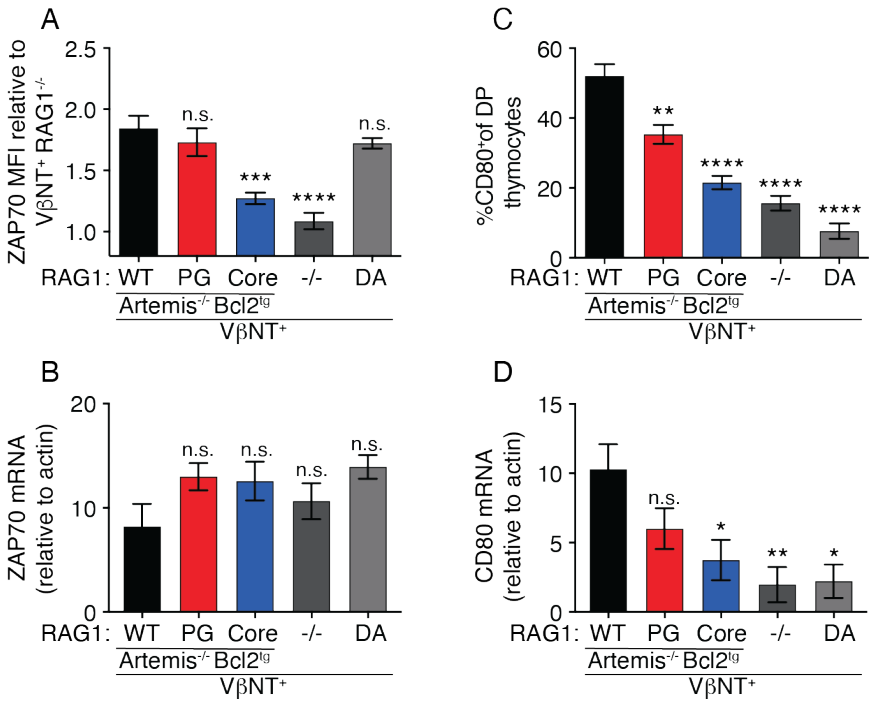


Figure 3-7: N-terminal region of RAG1 controls protein expression levels of signaling molecules via both transcriptional and non-transcriptional regulation

Mice described were engineered such that thymocytes progress to but get trapped at the double positive stage of T cell development (see Table 3-1). **(A)** Protein expression of ZAP70 measured by flow cytometry of trapped DP thymocytes ($CD4^+CD8^+CD69^-$) from respective mice. MFI reported normalized to $RAG1^{-/-}$ mice. Data combined from 3 independent experiments [N: WT=10, PG=3, Core=6, -/- =16, DA=3]. **(B)** ZAP70 mRNA expression from sorted DP thymocytes ($CD4^+CD8^+CD69^-$) from respective mice measured by qPCR. Data combined from 3 independent experiments [N: WT=8, PG=5, Core=7, -/- =12, DA=3]. **(C)** %CD80 of trapped DP thymocytes ($CD4^+CD8^+CD69^-$) from respective mice measured by flow cytometry. Data combined from 3 independent experiments [N: WT=14, PG=6, Core=9, -/- =14, DA=6]. **(D)** CD80 mRNA expression from sorted DP thymocytes ($CD4^+CD8^+CD69^-$) from respective mice measured by qPCR. Data combined from 3 independent experiments [N: WT=7, PG=5, Core=7, -/- =8, DA=3]. All data was analysed by one-way ANOVA and Dunnett's post test to compare each genotype individually with $V\beta NT^+RAG1^{WT}Artemis^{-/-}Bcl2^{tg}$ mice. Experiment performed by Charline Miot (Bassing Lab).

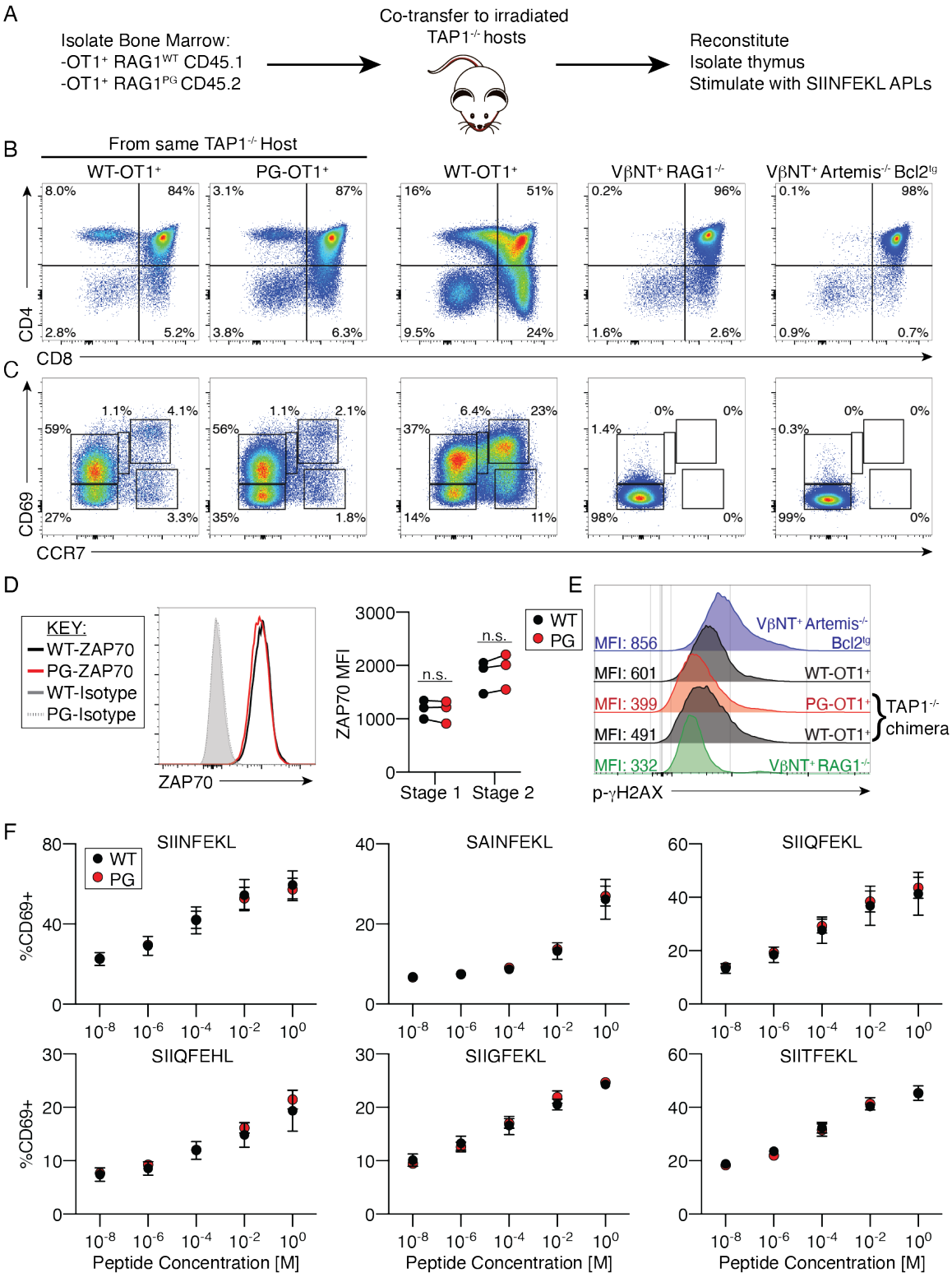

A
Isolate Bone Marrow:
-OT1+ RAG1WT CD45.1
-OT1+ RAG1PG CD45.2
Co-transfer to irradiated
TAP1-/- hosts
Reconstitute
Isolate thymus
Stimulate with SIINFEKL APLs
B
From same TAP1-/- Host
WT-OT1+
PG-OT1+
WT-OT1+
VβNT+ RAG1-/-
VβNT+ Artemis-/- Bcl2tg
CD4
CD8
C
CD69
CCR7
D
KEY:
WT-ZAP70
PG-ZAP70
WT-Isotype
PG-Isotype
ZAP70
ZAP70 MFI
n.s.
WT
PG
Stage 1
Stage 2
E
MFI: 856
VβNT+ Artemis-/- Bcl2tg
MFI: 601
WT-OT1+
MFI: 399
PG-OT1+
MFI: 491
WT-OT1+
TAP1-/- chimera
MFI: 332
VβNT+ RAG1-/-
p-γH2AX
F
SIINFEKL
SAINFEKL
SIIQFEKL
SIIQFEHL
SIIGFEKL
SIITFEKL
%CD69+
Peptide Concentration [M]

Figure 3-8: OT1^{+} TAP1$^{-/-}$ Chimeras do not provide a sufficient system to probe signaling differences between RAG1^{WT} and RAG1^{PG} thymocytes.

(A) Outline of experimental system. **(B-C)** Gross T cell development of RAG1^{WT}***or***^{PG}OT1^{+} thymocytes developing in TAP1$^{-/-}$ hosts compared to a normal RAG1^{WT}OT1^{+} thymus, and VβNT^{+}RAG1$^{-/-}$ or VβNT^{+}RAG1^{WT}Artemis$^{-/-}$Bcl2^{tg} thymuses as positive controls for developmental blocks at DP stage of thymocyte development. (B) pre-gating: live singlets. **(C)** pre-gating: live singlets, CD4^{+} *AND/OR* CD8^{+}, WT=CD45.1^{+} or PG=CD45.1^{-}. **(D)** ZAP70 expression in RAG1^{WT}***or***^{PG}OT1^{+} thymocytes developing in TAP1$^{-/-}$ hosts. Lines connect WT or PG thymocytes within the same TAP1$^{-/-}$ host. Analyzed by 2-way paired ANOVA with Sidak post-test. Representative of 3 independent experiments. **(E)** phospho-γH2AX staining from RAG1^{WT}***or***^{PG}OT1^{+} thymocytes developing in same TAP1$^{-/-}$ host, compared to a normal RAG1^{WT}OT1^{+} thymus, and VβNT^{+}RAG1$^{-/-}$ or VβNT^{+}RAG1^{WT}Artemis$^{-/-}$Bcl2^{tg} controls representative of at least 9 chimeric mice, and 3 controls. **(F)** Thymocytes from RAG1^{WT}***or***^{PG}OT1^{+}:TAP1$^{-/-}$ chimeric mice with incubated on peptide-loaded BMDCs for 2 hours, and CD69 expression of DP thymocytes was measured by flow cytometry. Peptides were canonical SIINFEKL peptide that the OT1 T cell receptor is specific for, and altered peptide ligands of various affinities. Peptides were incubated with BMDCs at respective concentrations for 2 hours prior to co-culture. Gating: Live singlets, thymocytes (excluding large BMDCs), CD4^{+}, CD8^{+}, CD45.1$^{+\text{ or }-}$.

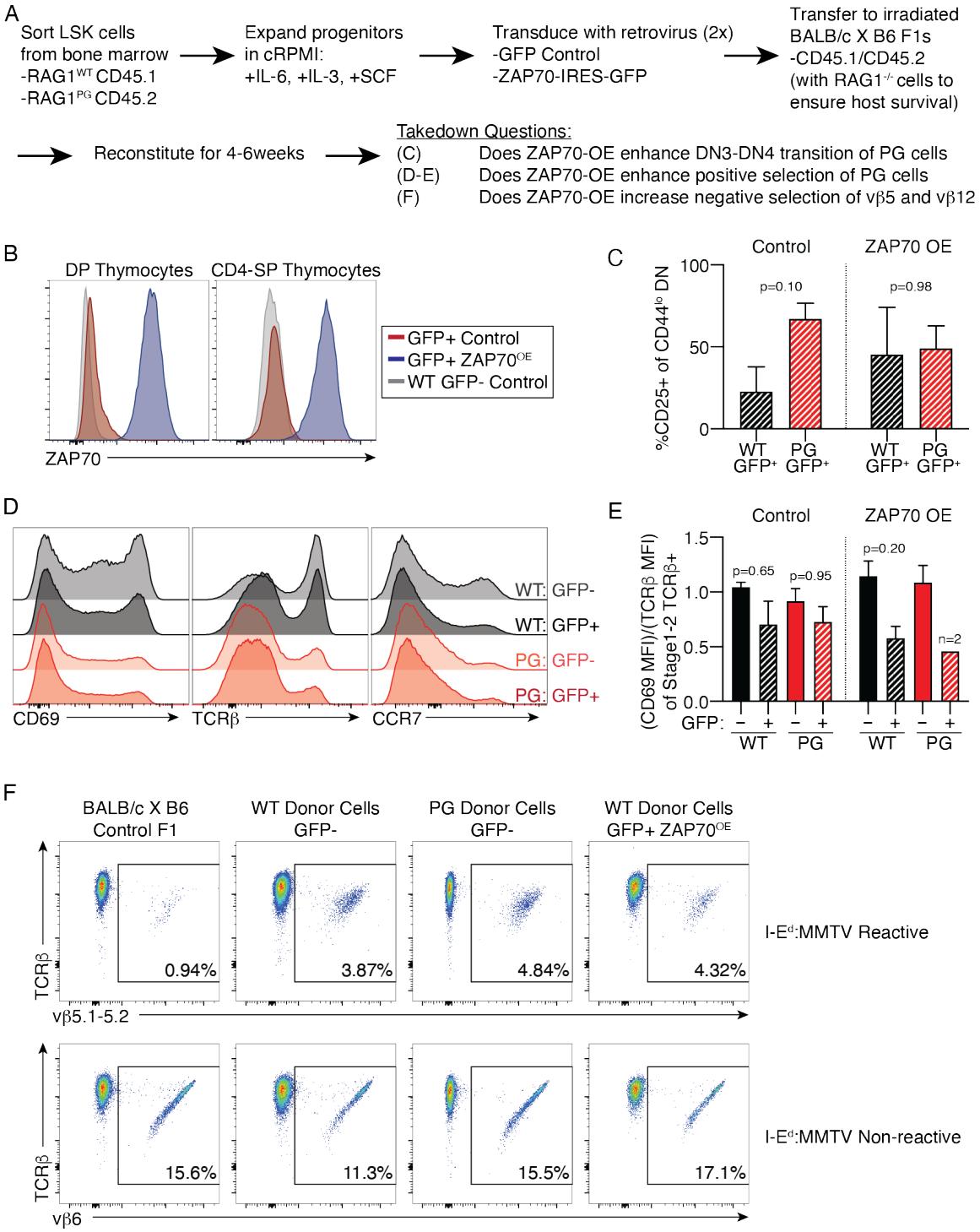


Figure 3-9: ZAP70 overexpression cannot rescue developmental and negative selection defects of PG thymocytes in an imperfect system

(A) Outline of experimental system. **(B)** ZAP70 expression in chimeric mice 5weeks post transplant. Gating: Live singlets, DP ($CD4^+CD8^+$); SP ($CD4^+CD8^-$); $GFP^{+\text{ or }-}$. **(C)** Investigation of DN3 to DN4

transition in chimeric mice. Showing %CD25+ of $CD44^-$ DN thymocytes as a proxy of the ratio between DN3 and DN4 thymocytes (gating: Live, singlets, $CD4^-$, $CD8^-$, $CD44^-$, GFP^{+or-} ; WT=$CD45.1^+CD45.2^-$, PG=$CD45.1^-CD45.2^+$). N.B. Host cells are $CD45.1^+CD45.2^+$ and are excluded from analysis. Data combined from 2 independent experiments. Data points were excluded if less than 1% GFP^+, or less than 1% from a respective donor. Analysed by 2-way ANOVA with Tukey post-test. **(D)** Investigation of positive selection in chimeric mice. CD69, TCRβ, and CCR7 expression is shown from a single chimeric mouse that includes all 4 populations of cells of interest: WT-GFP^-, WT-GFP^+($ZAP70^{OE}$), PG-GFP^-, PG-GFP^+($ZAP70^{OE}$). Gating: Live, singlets, $CD4^+$ *AND/OR* $CD8^+$, GFP^{+OR-} ; WT=$CD45.1^+CD45.2^-$, PG=$CD45.1^-CD45.2^+$). N.B. Host cells are $CD45.1^+CD45.2^+$ and are excluded from analysis. **(E)** Investigation of positive selection in chimeric mice. Showing ratio of CD69 MFI to TCRβ MFI in Stage 1 or Stage 2 thymocytes (gating: Live, singlets, $CD4^+$ *AND/OR* $CD8^+$, $CCR7^-$, $TCR\beta^+$; WT=$CD45.1^+CD45.2^-$, PG=$CD45.1^-CD45.2^+$). N.B. Host cells are $CD45.1^+CD45.2^+$ and are excluded from analysis. Analysed by 2-way ANOVA with Tukey post-test. **(F)** Investigation of negative selection via super-antigen mediated negative selection as described for Figure 3-3. The expression of $I\text{-}E^d$:MMTV reactive Vβ5.1-5.2 and unreactive Vβ6 TCRs are shown on Stage 4/5 thymocytes from chimeric mice, or a BALBc/B6 F1 control (control for efficient negative selection). Gating: Live, singlets, $CD4^+$ *AND/OR* $CD8^+$, $CCR7^+$ GFP^{-OR+}, WT=$CD45.1^+CD45.2^-$, PG=$CD45.1^-CD45.2^+$, BALBc/B6-F1=$CD45.1^+CD45.2^+$.

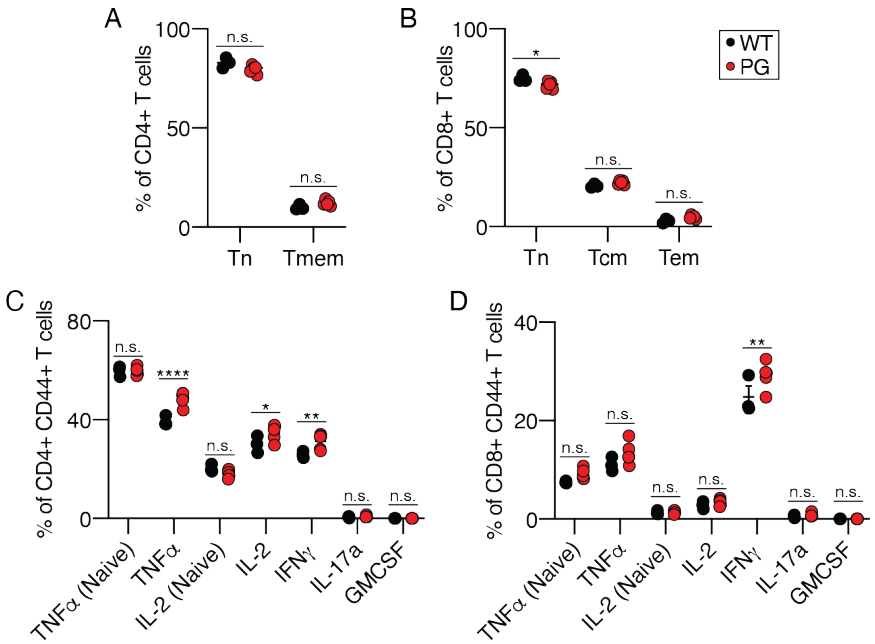


Figure 3-10: Peripheral PG T cells can make slightly more cytokines directly ex vivo compared to WT counterparts

Surface phenotype of CD4+ **(A)** and CD8+ **(B)** T cells from spleens of naïve WT or PG mice was measured by flow cytometry. Pre-gating: Live singlets, TCRβ+. Tn=Naïve T cells (CD44-CD62L+), Tmem=Memory T cells (CD44+), Tcm=Central Memory T cells (CD44+, CD62L+), Tem=Effector Memory T cells (CD44+ CD62L-). **(C-D)** T cells from naïve WT or PG mice were stimulated with PMA and Ionomycin + BFA/Monensin for 4 hours and intracellular TNFα, IL-2, IFNγ, IL-17A and GMCSF were measured by flow cytometry. Expression of CD44+ cells, unless specified by "Naïve" (CD44-). Gating: Live singlets, CD90.2+, CD4 *OR* CD8 (respectively). Analysed by 2-way ANOVA and Sidak post-test.

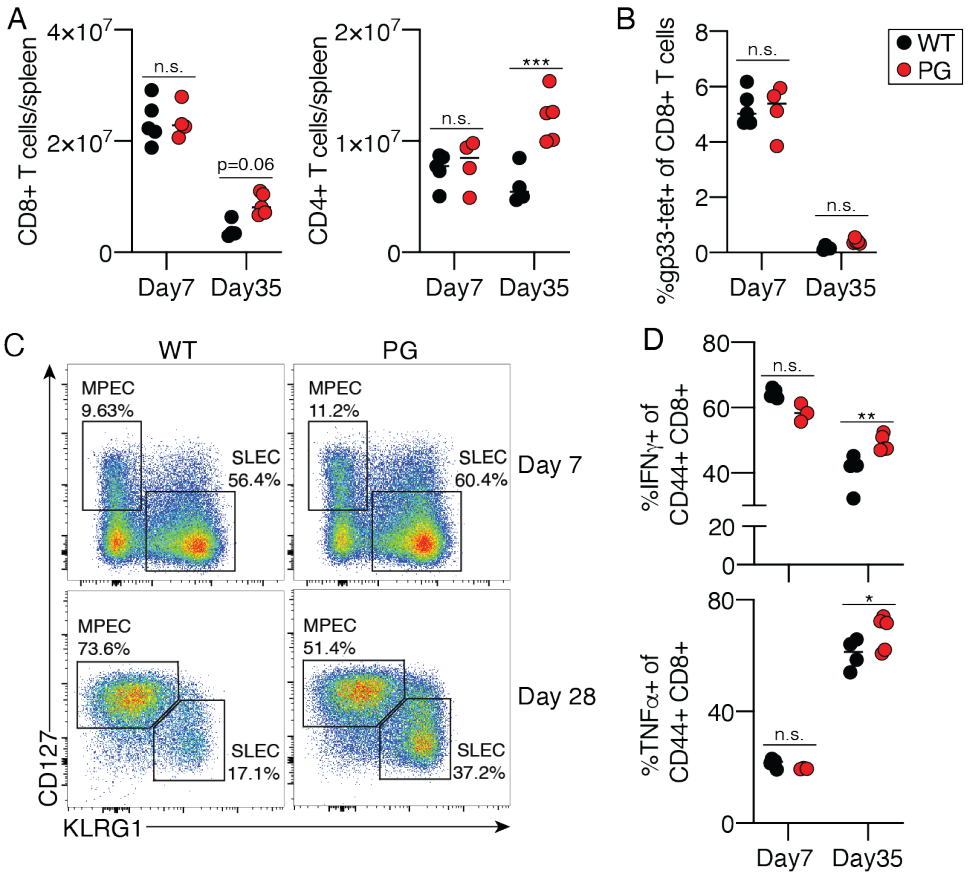


Figure 3-11: LCMV infection induces prolonged expansion of short lived, cytokine producing effector cells in PG mice

WT or PG mice infected with LCMV-armstrong and then sacrificed at day 7 (effector phase), or day 35 (memory phase) post-infection. CD8+ **(A)** or CD4+ **(B)** T cells were enumerated from spleens of infected mice. Gating: Live singlets, CD90.2+, CD4+ or CD8+ respectively. **(C)** Phenotype of CD8+ T cells from WT or PG mice at Day 7 (top) or Day 28 (bottom). MPEC=memory precursor effector cell. SLEC=short lived effector cell. Pre-gating: Live, singlets, CD90.2+, CD8+, CD44-. **(D)** IFNγ and TNFα from PMA and ionomycin stimulated CD8+ T cells isolated from WT or PG mice at Day 7 or Day 35 post-infection. Pre-gating: Live, singlets, CD90.2+, CD8+, CD44-. All statistical analyses were 2-way ANOVA with Sidak post-test.

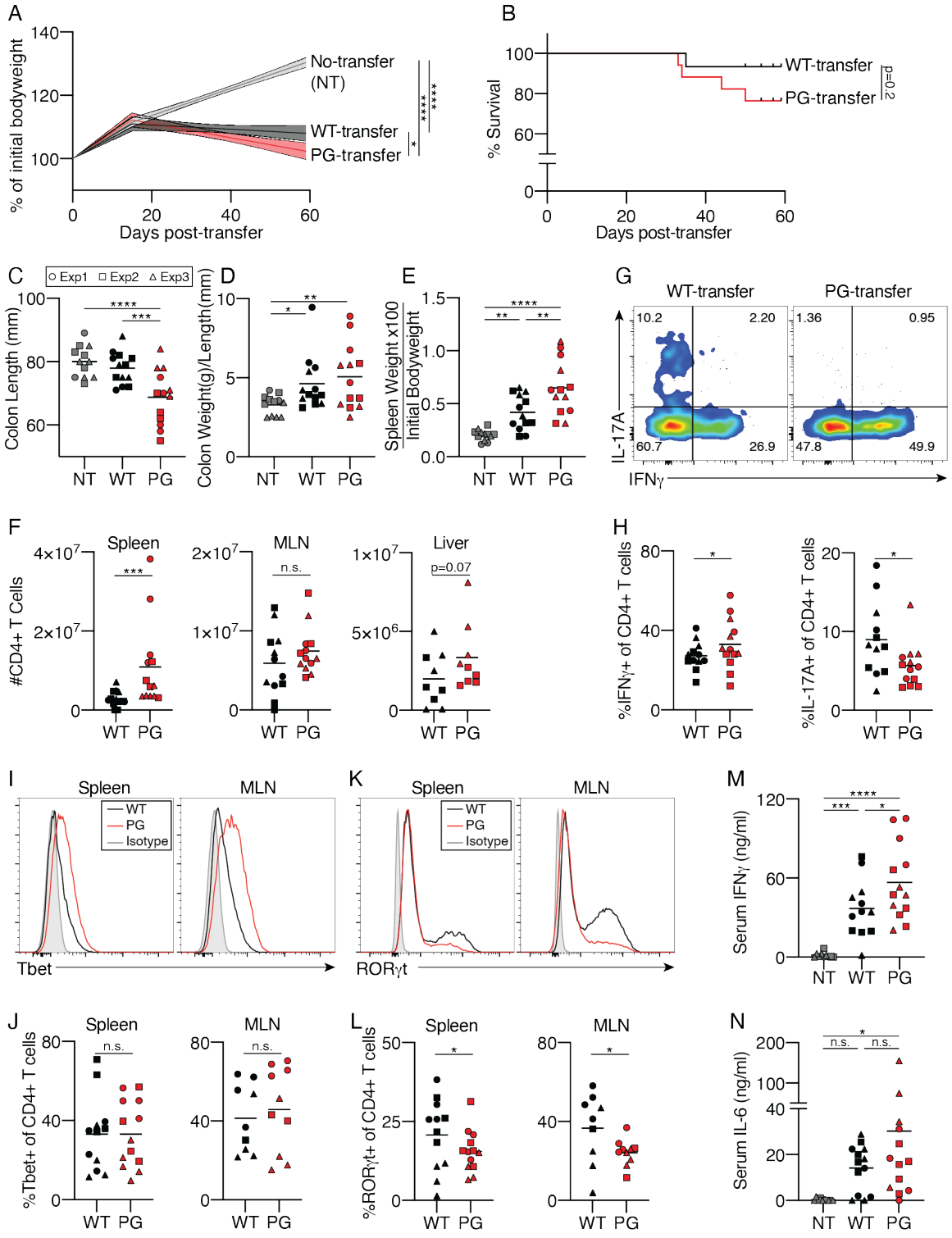


Figure 3-12: Naïve PG CD4+ T cells cause more severe colitis and disseminated inflammation when transferred into RAG1$^{-/-}$ hosts

Naïve CD4+ T cells were sorted from spleens of donor mice (CD90.2+, CD4+, CD8-, CD44-, CD62L+, CD25-) and 0.5E6 cells were transferred into co-housed 6 week old RAG1$^{-/-}$ hosts. The following analyses were performed. **(A)** Weights of RAG1$^{-/-}$ hosts that had no transfer of T cells (NT) or were transferred with WT or PG T cells. Data combined from 3 individual experiments (Total N: NT=13, WT=15, PG=17). Lines indicate fitted segmented-linear regression lines +/- 95% confidence interval. Day 15 was fixed as the inflection point. Difference between gradients of second slope was determined by linear regression analysis. **(B)** Survival curve of WT vs PG transferred mice. Data combined from 3-independent experiments. Statistical analysis = Log-rank Mantel-Cox test. Overall 0/13 NT mice died, 1/15 WT-transferred mice, and 4/17 PG-transferred mice. **(C)** Colon length of mice sacrificed at 8 weeks post-infection. **(D)** Colon weight divided by length as a measure of colon thickening. **(E)** Spleen weight normalized to mouse bodyweight pre-transfer of T cells. **(F)** CD4+ T cell counts in spleen, mesenteric lymph node (MLN) and liver. **(G)** Example staining of intracellular IL-17A and IFNγ from CD4+ T cells isolated from spleens of RAG1$^{-/-}$ mice transferred with WT or PG cells following restimulation with PMA and Ionomycin. Pre-gating: Live singlets, CD90.2+, CD4+, CD44+. **(H)** Combined IFNγ and IL-17A intracellular staining from 3-independent experiments. **(I)** Example Tbet staining from WT and PG cells from spleen or MLN from cells isolated 8weeks post-transfer. Pre-gating: Live singlets, CD90.2+, CD4+, CD44+. **(J)** Combined Tbet intracellular staining from 3-independent experiments. **(K)** Example RORγt staining from WT and PG cells from spleen or MLN from cells isolated 8weeks post-transfer. Pre-gating: Live singlets, CD90.2+, CD4+, CD44+. **(L)** Combined RORγt intracellular staining from 3-independent experiments. **(M-N)** Serum IFNγ(M) and IL-6(N) was analyzed by ELISA from mice 8 weeks post-transfer. (C-E, F, H, J, L, M, N) display data combined from 3-independent experiments. Experiment indicated by shape of data point (C). Horizontal bar indicates grand mean of combined data. All analysed by 2-way ANOVA with “Experiment” analysed as a variable. Statistics: Tukey post-test shown for all graphs with NT, WT, and PG data displayed. For graphs with WT vs PG only, statistic indicates the effect of genotype in 2-way ANOVA analysis.

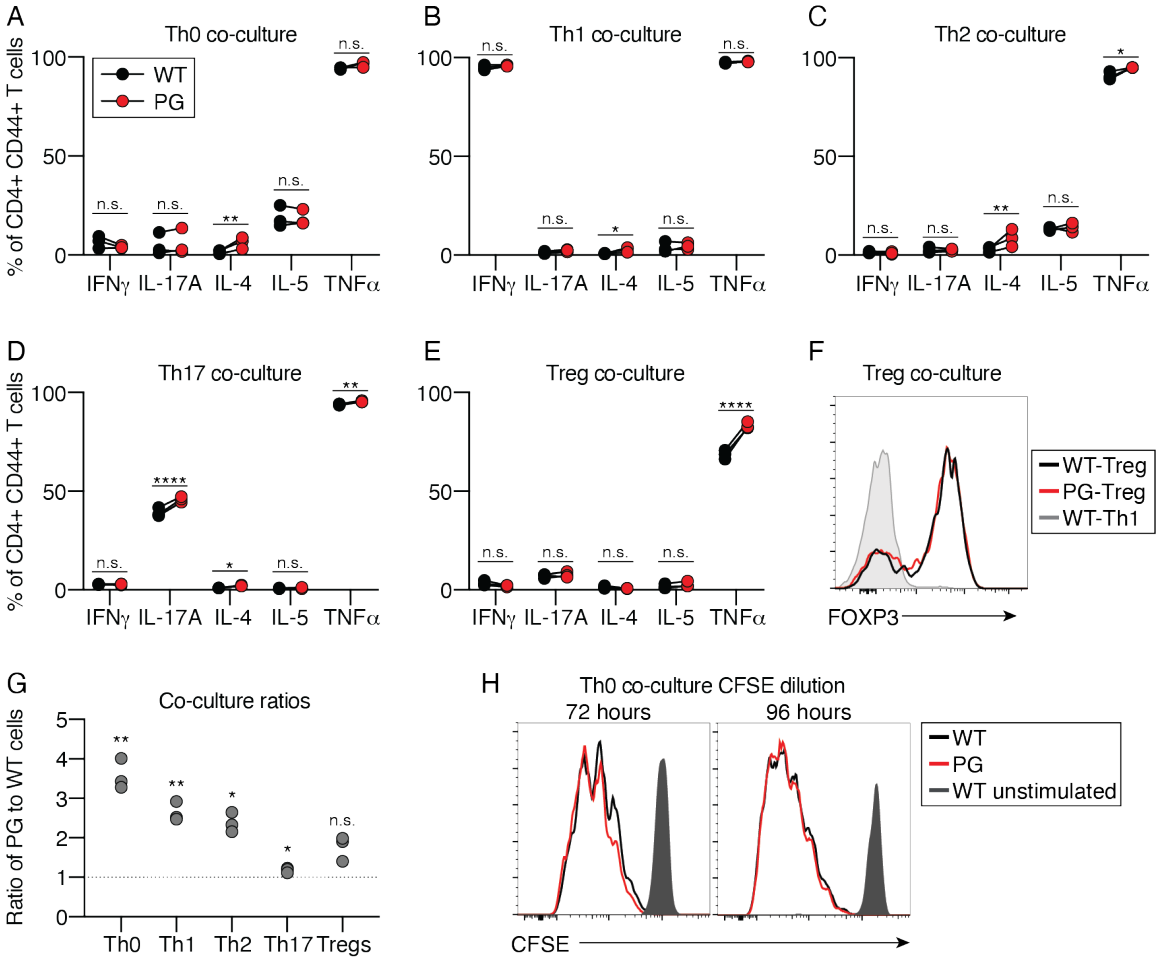


Figure 3-13: In vitro differentiated PG CD4+ T cells have slightly enhanced cytokine production and outgrow WT cells likely via enhanced survival

(A-G) Naïve (CD44- CD62L+ CD25-) CD4+ T cells were sorted from WT-CD45.1^{+} and PG-CD45.2^{+} mice. Cells were co-cultured in a 1:1 ratio of WT and PG cells in 96 well flat-bottomed plates with irradiated feeders, 10ug/ml anti-CD3, 5ug/ml anti-CD28, 30U/ml rh-IL2 and respective cytokines for T-helper skewing conditions. **(A)** Th0: IL-2 alone; **(B)** Th1: rmIL12, anti-IL4; **(C)** Th2: rmIL-4, anti-IFNγ, anti-IL12; **(D)** Th17: rmIL6, rmTGFβ1, anti-IL12, anti-IL4; **(E-F)** Treg: rmTGFβ1, anti-IL4, anti-IL12. Following 5 days of stimulation in respective conditions, cells were restimulated with PMA and ionomycin and stained intracellularly for IFNγ, IL-17A, IL-4, IL-5, and TNFα. Gating: Live singlets, CD4+, CD45.2^{-}(WT) *OR* CD45.2^{+}(PG). Lines between symbols indicated cells cultured within same well. Analysed by 2-way repeated measures ANOVA with Sidak post-test. **(F)** Intracellular FoxP3 stain of WT and PG cells co-cultured in Treg skewing conditions. WT-Th1 skewed cells used as negative control for FoxP3 staining. **(G)** Ratio of PG to WT cells within each co-culture condition at day 5. Statistics indicate 1-sample t-test comparing actual ratio to expected ratio of 1. **(H)** Naïve (CD44- CD62L+ CD25-) CD4+ T cells were sorted from WT-CD45.1^{+} and PG-CD45.2^{+}, stained with CFSE and co-cultured in Th0 conditions in a 1:1 ratio with anti-CD3/CD28 dynabeads for 72 or 96hrs. Gating: Live singlets, CD4+, CD44+, CD45.1^{+}(WT) *OR* CD45.1^{-}(PG)

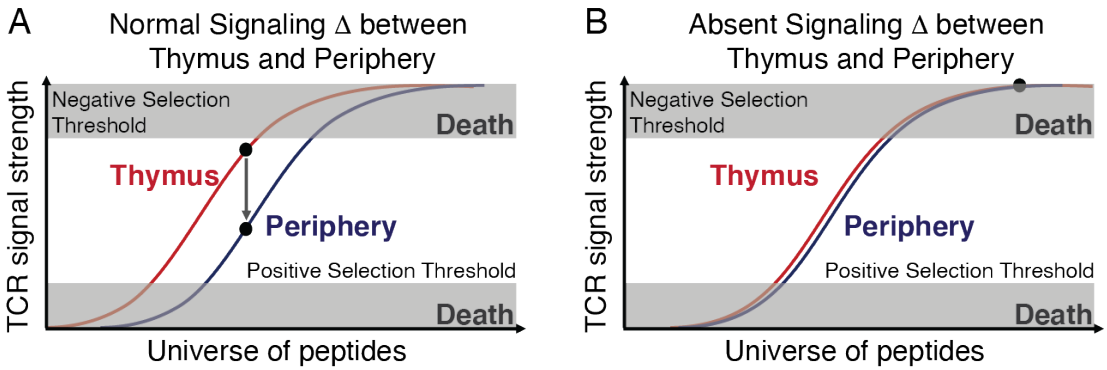


Figure 3-14: Model of signaling difference between T cells in the thymus vs. the periphery
(A) Normal setting where signaling delta exists between thymus and periphery. Example points to how a potentially self-reactive T cell that does not delete has a lower peripheral activation potential when thymic tuning is in place. **(B)** Setting where no signaling delta exists between thymus and periphery, for example in RAG1^{PG} mutant case. Here, peptides that are close to the deletion threshold will still be able to provide a strong signal in the periphery

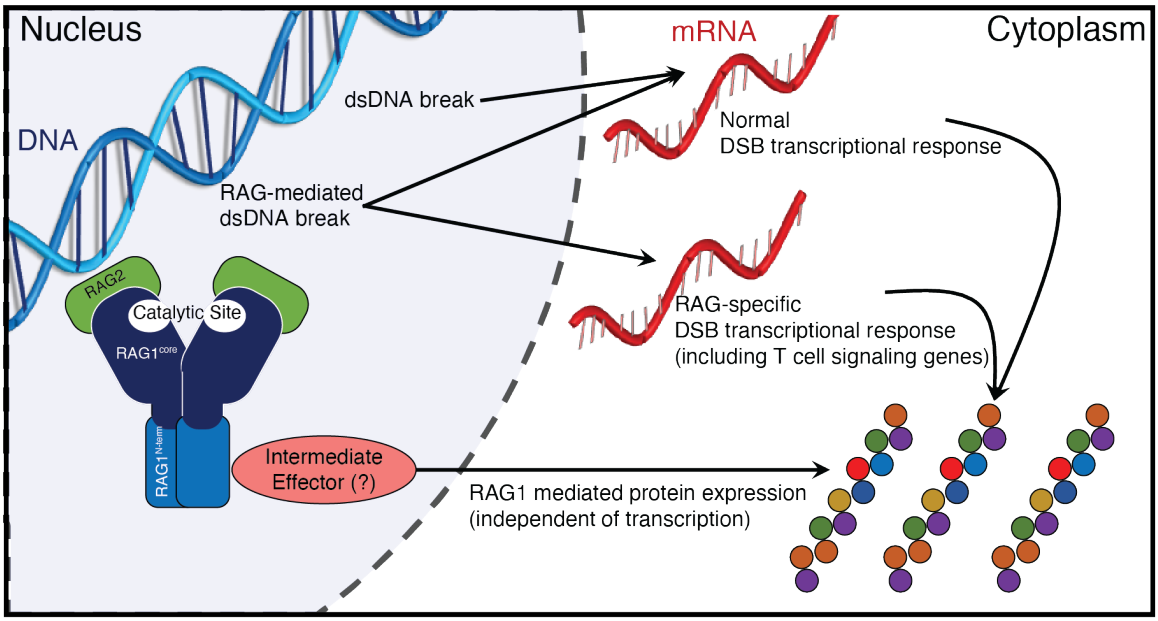


Figure 3-15: Model suggesting mechanisms as to how RAG1 contributes to upregulation of protein expression in the cytoplasm

It is known that dsDNA breaks initiate a transcriptional response, and that RAG-mediated DSBs initiate a RAG-specific transcriptional response (92). Additionally, we have shown that the RAG1 N-terminal domain, including the E3-ligase activity, is important in controlling protein expression through both transcriptional, and post-transcriptional regulation. Unless RAG-driven DSBs enhance transcription of a regulator of cytoplasmic protein expression (i.e. cytoplasmic E3 ligases, proteosome, etc.), because of the nuclear localisation of RAG1, RAG1 E3 ligase activity is unlikely to directly regulate cytoplasmic protein expression, and likely acts through an intermediate effector.

3.6: Tables

Table 3-1: Description of genetic background of mice used in Chapter 3

Gene	Genotype Description	Relevant Phenotypic Description
$RAG1^{NX}$	Q60X point mutation in RAG1 as described in detail in Chapter 2	N-terminal stop mutation that does not permit expression of full-length RAG1
$RAG1^{PG}$	P326G point mutation in RAG1 (53) as described in detail in Chapter 3	RAG1 ubiquitylation activity is inactivated
$RAG1^{Core}$	Mice only expressing core domain of RAG1 (47)	N-terminal region of RAG1 is absent. Impacts on T and B cell development (47, 48)
$RAG1^{-/-}$	RAG1 knockout (45)	Complete absence of RAG1 leads to absence of recombination and therefore B and T cell development
$RAG1^{DA}$	D708A point mutation in catalytic site of RAG1 (178)	RAG1 catalytic activity is absent.
$Artemis^{-/-}$	Knockout of DNA repair gene Artemis (179, 180)	DNA breaks are made by RAG complex but cannot be repaired. Developing lymphocytes are trapped at antigen receptor rearrangement steps
$Bcl2^{tg}$	Transgenic expression of survival factor Bcl2	Allows for survival of cells with DNA double stranded breaks.
VβNT	Transgenic expression of fully rearranged Vβ14 TCRβ allele (148)	Allows T cells to progress past DN3 transition even in the absence of recombination (i.e. $RAG1^{-/-}$ or $Artemis^{-/-}$)
Mouse Genetic Combination	**Relevant Phenotypic Description**	
$V\beta NT^{+}RAG1^{WT}$ $Artemis^{-/-}Bcl2^{tg}$	Thymocytes trapped at DP stage with continuous expression of $RAG1^{WT}$ but with inability to fix DSB breaks at TCRα locus, therefore preventing expression of a fully rearranged T cell receptor	
$V\beta NT^{+}RAG1^{PG}$ $Artemis^{-/-}Bcl2^{tg}$	As above, but with continuous expression of $RAG1^{PG}$	
$V\beta NT^{+}RAG1^{core}$ $Artemis^{-/-}$ $Bcl2^{tg}$	As above, but with continuous expression of $RAG1^{core}$	
$V\beta NT^{+}RAG1^{-/-}$	Thymocytes trapped at DP stage without RAG1 expression therefore no DSBs and inability to rearrange TCRα locus, therefore preventing expression of a fully rearranged T cell receptor	
$V\beta NT^{+}RAG1^{DA}$	Thymocytes trapped at DP stage with continuous expression of $RAG1^{DA}$, therefore no DSBs and inability to rearrange TCRα locus, therefore preventing expression of a fully rearranged T cell receptor	

Table 3-2: Gross non-colitis disease phenotypes observed in mice transferred with Naïve CD4$^+$ T cells from WT or PG donors

Disease Phenotype	WT-transferred mice	PG-transferred mice
Premature death	1/15	4/17
Severe dermatitis	0/15	3/17
Chylous ascites	0/15	1/17
Severe ataxia	0/15	1/17
Severe small intestine enlargement	0/15	1/17
Total Adverse Events	1	10

CHAPTER 4: DISCUSSION

4.1: Overview

The findings described in this thesis can be broadly split into two related areas of RAG biology. In Chapter 2 I describe internal translation of RAG1 that leads to the expression of N-terminally truncated RAG1 isoforms. I describe how a mouse with an N-terminal nonsense mutation does not present with SCID as may be expected, but rather N-terminally truncated RAG1 isoforms in the absence of full-length RAG1 can lead to impaired VDJ recombination and lymphocyte development. In Chapter 3 I describe how the N-terminal region of RAG1, and specifically E3-ligase activity of RAG1, can influence cytoplasmic protein expression in developing T cells. Additionally, I show that positive and negative selection are impaired in mice lacking RAG1 E3 ligase activity, and that peripheral T cells from these mice are hyperactive and can cause more widespread pathology when transferred into $RAG1^{-/-}$ hosts.

The purpose of this discussion is to expand upon a number of birds-eye-view ideas that have arisen during these investigations. In section 4.2 I discuss what the normal expression of N-truncated RAG1 isoforms may mean for developing T cells. In section 4.3 I propose a number of experiments to determine the mechanism by which RAG1 internal translation occurs. In section 4.4 I advance the idea of internal translation of eukaryotic proteins outside of RAG, and query whether enough attention is paid to this phenomenon. In section 4.5 I propose two mechanisms by which RAG1 may operate to enhance signaling in thymocytes, and thus positive and negative selection, and the future directions in which I hope these studies will head. Finally, in section 4.6 I briefly describe the true

trajectory of this story, and how the true pathway to my coming to study RAG1 became one of the most important lessons of my Ph.D.

4.2: RAG1 internal translation: implications for lymphocyte development

The most important finding in the description of the $RAG1^{NX}$ mice is that a 5' nonsense mutation does not lead to loss of functional RAG1 protein. This raises a number of important conceptual points. First, that RAG1 has many naturally occurring, N-truncated isoforms is of considerable interest. RAG1 has been studied for more than 30 years as a single protein, yet it is intriguing to consider that naturally occurring smaller RAG1 isoforms may have important roles in normal RAG function. Already we know that the N-terminus of RAG1 has important functions in aiding recombination activity. For example, the RAG1 N-terminus contains an E3 ubiquitin ligase that modifies histones to enhance RAG cleavage (49, 52), and regions that bind the VprBP kinase to help repair RAG DSBs (54) (Figure 1-1). Translation from M1 would retain the RAG1 RING domain but may ablate VprBP binding (Figure 4-1 B), while translation from M2 would likely ablate RING domain activity. Translation from M4 and M5 would result in loss of portions of the $RAG1^{core}$ domain and will likely have no recombination activity (Figure 4-1 C). Whether these internally translated proteins are present at biologically significant levels is of considerable interest, since smaller RAG1 isoforms could form part of the $RAG1_2/RAG2_2$ (RAG) heterotetrametric complex and be important for the regulation of normal function.

In order to test whether these RAG1 isoforms are expressed at high enough levels to influence T cell development and RAG1 function, the Bassing/Behrens Lab collaboration has planned to create a mouse lacking the ability to form internally translated protein isoforms. Using the recently developed Easi-CRISPR system (181, 182), which

allows replacement of up to ~2kb of DNA, mutation of the methionines at M1, M2, M4, and M5, to isoleucines can be achieved with one CRISPR reaction. We expect that the resulting mouse should only be able to express the full-length RAG1 isoform (Figure 4-1 A). I hypothesize the following possible outcomes upon examination of this mouse, which I will term the $RAG1^{FL}$ mouse in the following discussion to denote the presence of only full-length(FL) RAG1:

- No difference between $RAG1^{WT}$ and $RAG1^{FL}$ mice. Smaller isoforms are not expressed to a high enough level in $RAG1^{WT}$ mice to influence the activity of RAG1.
- RAG1 internal translation isoforms (especially the M4/M5 isoforms) negatively inhibit RAG1-FL VDJ recombination such that $RAG1^{FL}$ mice have enhanced recombination efficiency. As such, I would expect that developing T cells will progress through TCRβ and TCRα selection stages more rapidly. Additionally, I may expect to see the appearance of more dual TCR expressing cells. While the dogma for many years was that all T cells express a single receptor, this is not true 100% of the time. 1-3% of T cells express two in-frame TCRβ alleles (62, 148, 183), and ~30% of T cells express two in-frame TCRα alleles (63, 184, 185). A number of hypotheses have been posed as to how allelic exclusion is regulated/enforced (143). One such mechanism is the asynchronous recombination model, which states simply that the VDJ reaction is of such low efficiency, it is unlikely that both alleles are rearranged together. This allows for a TCR signal (either pre-TCR at TCRβ rearrangement, or TCR$\alpha\beta$ signaling/positive selection at TCRα rearrangement) to downregulate RAG1/RAG2 expression and prevent further rearrangement of the other allele (62). Much of this asynchrony is

postulated to be due to difficult chromatin landscapes that limit RAG accessibility (128, 129). It is possible that having RAG complexes containing N-truncated RAG1 isoforms that we have shown have inefficient recombination activity is another mechanism by which inefficient VDJ recombination is achieved, and thus allelic exclusion is enforced. This is of further interest since the only known role of the RAG1 E3 ligase activity is to ubiquitylate histones to enhance RAG accessibility, and chromatin accessibility is the prime mediator of the asynchronous recombination model. Therefore, I would hypothesize that the $RAG1^{FL}$ mouse would have a higher frequency of allelic inclusion (dual TCRα and/or TCRβ expressing T cells), if N-truncated isoforms inhibit VDJ recombination.

- The findings described within Chapter 3 of this thesis suggest that the N-terminal domain, and specifically the E3 ligase activity of RAG1, is necessary for efficient positive and negative selection. Therefore, I hypothesize that the $RAG1^{FL}$ mouse will have enhanced positive and negative selection (as measured by the assays in Figure 3-3). Likewise, the expression of signaling molecules such as ZAP70 and SYK would be increased in the absence of N-truncated RAG1 isoforms. It is possible that positive and negative selection are at maximum efficiency in $RAG1^{WT}$ mice already, and the potentially more interesting effects on selection may be observed within agonist-selected T cells. A specialized group of $\alpha\beta$ T cell subtypes including natural Tregs (nTregs), invariant natural killer T cells (iNKT cells), and CD8$\alpha\alpha$ intraepithelial lymphocytes (CD8$\alpha\alpha$ IELs), are created by a mechanism known as “agonist selection” (82-84). These T cells have higher affinities towards self-antigens, yet why these T cells are not directed for death via negative selection

is not well known. It is attractive to speculate that self-reactivity of these developing T cells is not detected as TCR signaling in these cells is not tuned up as efficiently, because of increased expression (relative to 'normal' αβ T cells) of N-truncated RAG1 isoforms. Therefore, I would hypothesize that $RAG1^{FL}$ mice would have reduced development of Tregs, iNKT cells, and CD8αα IELs, as these cells would more likely be directed towards death.

Mentioned in this last point is an allusion to differential expression of internally translated isoforms. Whether there are different ratios of N-truncated RAG1 isoforms in different thymocytes is as of now unclear, since all protein expression analysis has been on bulk cells. Additionally, since RAG1 expression is rapidly downregulated following β-selection and positive selection, it may be nearly impossible to calculate cell to cell differences, especially since the effects of RAG1 are detected after RAG1 expression has ceased. One possible method, aside from removing the ability to make N-terminally truncated isoforms (as discussed in this section), is to understand how expression of these isoforms is regulated, and target effectors driving this mechanism.

4.3: RAG1 internal translation: mechanism and future directions

While I have shown in Chapter 2 that internal methionines are necessary for the formation of internally translated RAG1 isoforms, complete mechanism has not been described in this thesis. In fact, while examples of such proteins are increasing, mechanisms driving internal translation have not been well characterized for eukaryotic proteins. Such mechanisms are more commonly adopted by viruses to encode multiple proteins within limited genetic material (186). These mechanisms include utilization of internal ribosome

entry sites (IRESs), leaky scanning (i.e. bypassing of translation initiation sites), or ribosome shunting (186), described below.

Most eukaryotic proteins are translated via CAP-dependent mechanisms, whereby the mRNA 5'-m_7G cap (present at the 5' end of all eukaryotic mRNAs) is bound to the 'cap-binding complex', a protein complex that then recruits the 40S subunit of the ribosome (187, 188). The 40S subunit then scans the mRNA from 5' to 3' until it finds an AUG surrounded by a favorable Kozak sequence, which the remainder of the ribosome forms around (i.e. the translation initiation site/TIS), and translation begins (187, 188). Leaky scanning is simply the act of a ribosome bypassing translation initiation sites before initiating translation off 3' TISs. Such a mechanism has been described for the expression of the eukaryotic protein Pim2 (113), and is CAP-dependent. Shunting is a mechanism that involves the secondary structure of the mRNA forming a hairpin which can prevent ribosome movement. The ribosome then is transferred past the hairpin to internal TISs, and has been described for Hsp70 mRNA (114). Finally, IRES-directed translation, a CAP-independent mechanism, which allows ribosome assembly within central regions of the mRNA. This can allow for translation to begin downstream of the most 5' TIS. Approximately 30 mammalian mRNAs have IRESs, and initiate translation at an internal TISs (105-112).

Given the paucity of examples of shunting and leaky scanning for eukaryotic proteins, the most attractive hypothesis is that RAG1 contains an IRES between the annotated TIS, and M1. In fact, ~10% of eukaryotic mRNAs are hypothesized to contain IRESs, although many of these appear in 5' untranslated regions and would not theoretically initiate internal translation initiation (189). There is very little sequence

similarity between IRESs, as evidenced by the striking lack of IRES sequence similarity even within the Myc family of IRESs (190). This makes computational-based IRES prediction difficult, and experimental validation essential. In order to test whether RAG1 contains an IRES, I have proposed to introduce a hairpin into the RAG1 mRNA, 5' of the canonical TIS to halt ribosome scanning. The size of the inserted mRNA hairpin directly correlates with the ability for scanning-based translation to occur (191, 192). If RAG1 N-truncated isoforms are a result of ribosome scanning (or shunting), then the addition of an RNA hairpin should prevent expression of both full-length, and N-truncated RAG1 isoforms. However, the presence of an IRES in RAG1 would be indicated by absence of full-length RAG1, but unaffected expression of N-truncated isoforms.

IRESs are somewhat complicated RNA secondary structures (186). One of the more surprising facets of the Q60X mutation (in the $RAG1^{NX}$ mouse), is that nonsense-mediated degradation does not occur to a sufficient degree such that no RAG1 isoforms are made. RNA binding proteins often bind secondary structures within mRNA, including IRESs, and often proteins that bind mRNA are able to enhance mRNA stability (193). One such protein that is known to bind IRESs of a number of eukaryotic mRNAs and enhances mRNA stability is the polypyrimidine tract-binding protein (PTB) (193, 194). It is attractive to speculate that RAG1 does not undergo nonsense-mediated degradation to a significant degree due to the presence of RNA binding proteins, such as PTB, which stabilize RAG1 mRNA even in the presence of an early nonsense mutation. Importantly, I have not shown that $RAG1^{Q60X}$ mRNA is expressed as efficiently as $RAG1^{WT}$ RNA which is an immediate future direction.

Finally, IRES-mediated translation from eukaryotic mRNAs has been shown to be activated during periods of cellular stress (104, 189, 195). RAG1, by virtue of its own DNA-cutting activity, is present during a state of remarkable stress. It would be interesting to see whether the expression of RAG1 N-truncated isoforms increases after the initiation of the first RAG-mediated DNA break, and thus activation of the DSB response. A simple first experiment would be to compare expression of N-truncated isoforms in $RAG1^{DA}$ vs $RAG1^{WT}$ $Artemis^{-/-}$ thymocytes, with the expectation that the DSBs in $RAG1^{WT}$ $Artemis^{-/-}$ mice may activate internal translation from any putative RAG1-IRES.

4.4: Internal translation initiation: rethinking the interpretation of nonsense and frameshift mutations

That alternative translation initiation may be a natural phenomenon occurring widely within eukaryotic cells is still an underappreciated occurrence and may have many important consequences. Ribosome profiling in eukaryotic cells suggests that internal translation initiation generates N-truncated isoforms for ~15% of the ~20,000 human proteins (102, 103, 196), yet our understanding of mechanisms that control internal translation initiation and how resulting N-truncated isoforms function normally and contribute to disease are at infancy. This description of alternative translation of RAG1 contributes to the expanding database of alternatively translated proteins, and the exact mechanism of translation of RAG1 alternative isoforms is a subject of future investigation as described above.

A global consequence of internal translation initiation outside of RAG1 is that nonsense mutations, or frameshift mutations leading to early translation termination may not always result in the absence of a protein. Indeed, it is attractive to speculate that internal translation initiation sites may exist as a mechanism by which to protect from

deleterious nonsense or frame-shift mutations, as is the case with the RAG1^{NX} mice that do not have the expected complete-SCID phenotype. In fact, it has been estimated that the average human genome contains ~100 putative "loss-of-function" (LoF) variants, yet the vast majority of these are not disease causing (197). 70% of the putative LoF variants were either nonsense or frameshift mutations (indels) (197), however, it is possible that internal translation from alternative TISs can compensate for, and lead to expression of functional proteins despite putative LoF mutations.

The potential for internal translation initiation is important to consider during analysis of clinical exome/genome sequencing as interpretations can vary drastically. Frameshift or nonsense mutations in one allele of a gene are often deemed unlikely to be disease causing if a wild type allele is still present. Alternatively, pathogenesis caused by a monoallelic frameshift or nonsense mutation is assumed to be a result of haploinsufficiency, and homozygous mutations presume LoF (198). However, as an example, N-terminal truncations of certain proteins due to early frameshift/nonsense mutations and internal translation, may delete N-terminal auto-regulatory domains, resulting in constitutively active proteins rather than complete absence. Our lab has identified an example of such a protein, described in Box 1 below.

Box 1: Internal translation of NLRP1 in a VEO-IBD patient

NLR (nucleotide-binding domain and leucine-rich-repeat containing) family members are cytosolic proteins that are key parts of the innate immune response poised to recognize microbial products that gain entry to the cytoplasm (199). NLRP1 is one family member (of an estimated 20 in the human genome) that is conserved across most mammalian

species (199). Human NLRP1 contains in the N-terminal region, an autoinhibitory domain (PYD) that self-oligomerizes with the C-terminal CARD domain (200). Specific activators of human NLRP1 are not well characterized, however, auto-catalytic cleavage at the FIIND domain between the PYD and CARD domains releases the NLRP1 CARD domain to then oligomerize with ASC and induce formation of the "inflammasome" (199, 201). The activated inflammasome cleaves caspase 1 into its active form, which leads to proteolytic cleavage of the inflammatory cytokines, pro-IL-1β and pro-IL-18, into active forms (201).

In a cohort of whole exome sequences (WESs) from very early onset inflammatory bowel disease (VEO-IBD) patients, we identified one patient with a heterozygous frameshift mutation in the N-terminal region of NLRP1 caused by a four-nucleotide deletion (c.158_161delTGGC). This variant was classified as a "Variant of uncertain significance" by XomeDX WES reporting. The patient presented with VEO-IBD, Crohn's Disease, and cellulitis. Given the position of the mutation (very N-terminal), the fact that NLRP1 contains an N-terminal autoinhibitory domain, and that IBD/Crohn's disease is commonly associated with high IL-1β/IL-18 levels (202-204), we were interested in determining whether this mutation could potentially be disease causing.

I hypothesized that NLRP1 may be translated from downstream internal TISs, and that since the resulting N-truncated NLRP1 products would lack the autocatalytic domain, spontaneous NLRP1 cleavage and inflammasome activation may occur. While this project is still in its infancy, I have now shown (using similar *in vitro* methods to that of RAG1 in Chapter 2) that despite the presence of the dTGGC frameshift mutation, N-truncated NLRP1 isoforms are still expressed. Whether these isoforms are hyperactive is the subject

of ongoing experimentation. This finding alone, however, accentuates the importance of the awareness around alternative outcomes of nonsense and frameshift mutations.

Finally, internal translation initiation is also an important factor to consider when designing conditional knockout mice using the cre-lox system, where typically exons are flanked by loxP sites and excised following cre expression leading to frameshifted proteins (205). While the assumption is that this frameshifted mRNA will be degraded by nonsense-mediated-decay, it is possible that translation initiation downstream of the excised region may result in protein expression. This is in fact an issue that has recently been described in a cre-lox mouse model used by the Wherry Lab at the University of Pennsylvania (Unpublished, Personal Communication, Omar Khan-Wherry Lab).

4.5: RAG1 as a biological clock to tune TCR signaling: mechanism and future directions

In Chapter 3 I discussed extensively the proposition that RAG1, specifically through its N-terminal domain and E3-ligase activity, acts as a biological clock to alter the proteome of developing thymocytes in order to enhance TCR signaling, and positive and negative selection. In this section I wish to introduce some more speculative hypotheses as to mechanisms that may drive such selection, and future experiments that are proposed to dissect these mechanisms.

Briefly to summarize: RAG-mediated DSBs have been shown to induce a transcriptional program that includes the expression of many lymphocyte-specific genes (92-94). In addition, the Behrens and Bassing labs have shown that the N-terminal domain of RAG1 is important for the post-transcriptional regulation of a number of proteins

including Pim2, SYK, and ZAP70. I have hypothesized that increased expression of signaling proteins during T cell development influences thymocyte-specific TCR tuning, to enhance thymocyte signaling, and improve positive and negative selection (I reiterate here that I have not yet linked decreased ZAP70 or SYK expression in thymocytes with reduced TCR signaling – attempts to elucidate differences in signaling between $RAG1^{WT}$ and $RAG1^{PG}$ mice are ongoing). Once thymocytes have completed development in the thymus, TCR sensitivity is reduced as the effects of RAG on proteins such as ZAP70 are not permanent, and T cells that were positively selected upon self-antigen in the thymus can no longer respond to that same self-antigen in the periphery.

Within this model exist a number of gaps. Notably is the absence of clear mechanism behind the post-transcriptional control of such proteins. One conceptual difficulty is the fact that RAG1 is strictly localized to the nucleus (206, 207), while proteins such as ZAP70 and SYK are cytoplasmic and associate with receptors at the plasma membrane. Therefore, it is unlikely that there is direct contact between the RAG1 N-terminal domain with these proteins, and thus the attractive hypothesis that RAG1 directly ubiquitylates such proteins is unfeasible. In light of this I propose two, non-mutually exclusive, hypotheses:

1. Transcriptional activation of a set of genes is achieved via the known activity of the RAG1 E3 ligase activity; i.e. ubiquitylation of histones. (Figure 4-2 A) This ubiquitylation enhances RAG access to recombination sites (52), yet it is known that RAG binds many parts of the genome (208) and could theoretically enhance chromatin access at a number of genes via similar a mechanism. Interestingly, the N-terminal region of RAG1 is required for localization of RAG1 to H3K27Ac histone

marks at enhancers (208) (Figure 4-2 A). Such genes may encode proteins that have known abilities to control protein expression levels to enhance protein translation (e.g. translational elongation factors, RNA binding proteins that stabilize transcripts, micro-RNAs that stabilize transcripts), or that slow rate of protein turnover/enhance protein stability (e.g. E3 ligases, deubiquitylating enzymes (DUBs), repression of proteasome). Experiments to test whether translation vs. protein stability is the major contributor to enhanced protein expression in the presence of RAG1 N-terminal domains is actively underway in the Bassing Lab, with early evidence (for Pim2) suggesting protein stability is enhanced by the RAG1 N-terminal domain.

2. The RAG1 E3 ligase activity is necessary for the ubiquitylation of an intermediate effector that regulates expression of cytoplasmic proteins. Such intermediate targets may include other E3 ligases that have activities within the nucleus and cytoplasm, such as Smurf1 (209); or transcription factors that drive expression of other protein control mechanisms as described in point 1 (Figure 4-2 B).

Potential mechanisms that are less attractive include RAG1 mediated alternative splicing mediated by known interactions between splice factors and the RAG1 N-terminal domain (57) (Figure 4-2 C). Or, RAG1 mediated stabilization of mRNA, or enhanced translation via the RAG1 RNA binding domain (29) (Figure 4-2 D).

In order to investigate these possibilities, the Bassing and Behrens labs have planned a number of unbiased experiments to measure differences between transcriptomes (RNA-Sequencing), whole cell proteomes (whole cell proteome mass-spectrometry), and ubiquitylated proteomes (KεGG mass spectrometry. Briefly: trypsin

digestion of ubiquitin leaves behind diglycine residues, i.e. KεGGs. Immunoprecipitation with an anti-KεGG antibody followed by mass spectrometry identifies ubiquitylated proteins (210-212)). These experiments will be performed using DP thymocytes expressing $RAG1^{WT}$, $RAG1^{core}$, $RAG1^{PG}$, or $RAG1^{DA}$ from the mice described in Table 3-1 and Figure 3-7. These experiments serve two purposes: First, it is unlikely that ZAP70, SYK, and CD80, are the only proteins regulated by the presence of RAG1 N-terminal domains, and these methods will identify other examples; and second, we hope to identify network hubs via protein interactome analyses and transcriptional pathway analysis that can better point to mechanisms behind the extra-recombination activities of RAG1.

4.6: The true story of this thesis

The road to RAG1 was as far from planned as I can imagine. The goal of this section, however, is not to deliver a self-pitying soliloquy, but is to briefly describe some of the most important scientific lessons I have learned through what has been somewhat of an unorthodox road to a Ph.D. thesis.

I joined the Behrens Lab to study the role of the cytokine IFNγ in the pathogenesis of the hyper-inflammatory disorder, familial hemophagocytic lymphohistiocytosis (HLH). HLH is a disease driven by the hyperactivation of CD8+ T cells that lack the ability to kill infected target cells, and make uncontrolled levels of IFNγ resulting in off-target pathology (213, 214). Mutations in genes that are required for efficient CD8+ T cell mediated cytotoxicity, including perforin, are common causes of disease in humans (213). As such, the Behrens lab has utilized perforin deficient mice to study mechanisms of FHL pathogenesis (215-217). IFNγ was believed to be the key cytokine involved in driving HLH,

and clinical trials led to the approval of an IFNγ-blocking monoclonal antibody for HLH treatment. However, there were also reports of patients presenting with HLH but who were unable to respond to IFNγ-signaling due to mutations in the IFNγ-receptor (218). Therefore, my proposed thesis project was to investigate the necessity of IFNγ-signaling in driving HLH-pathogenesis by utilizing Perforin/IFNγ ($Prf1^{-/-}IFN\gamma^{-/-}$) double knockout mice (which we published in 2019 (219)).

However, 6-months into the project, I showed that the $Prf1^{-/-}IFN\gamma^{-/-}$ mice had thymi with only a third of the cellularity of normal WT mice, and a defect in DN3 to DN4 progression. *Lesson 1: characterize T cell development, B cell development, and baseline populations of immune cells in naïve mice of a novel genotype.* It took a long time, and many seemingly worthless experiments to decide that it was unlikely that perforin, IFNγ, or an interaction between the both of them, was causing this defect in T cell development. Finally, because we wanted to utilize the $Prf1^{-/-}IFN\gamma^{-/-}$ mice for the purposes of the project above, but we did not know what other gene had been impacted, and therefore could not genotype these mice effectively, we performed whole exome sequencing (WES).

WES revealed a homozygous early nonsense mutation in RAG1 (Q60X). These $RAG1^{NX}$ mice were then backcrossed to WT C57BL/6 mice to obtain $Prf1^{+/+}IFN\gamma^{+/+}RAG1^{NX/NX}$ mice. Similarly, the $Prf1^{-/-}IFN\gamma^{-/-}$ mice were re-bred, and over the course of 18-24 months, one project had become two. *Lesson 2: Mice are not clones. Be conscious that spontaneous mutations occur, and many are likely present in all strains of mice.* Since Prf1 and IFNγ are on the same chromosome, in order to obtain $Prf1^{-/-}IFN\gamma^{-/-}$ mice a rare chromosomal cross-over event is required. Therefore, the entire line of $Prf1^{-/-}IFN\gamma^{-/-}$ mice had been derived from a small pool of founders which allowed for homozygous

acquisition and a phenotypic effect due to the Q60X mutation (heterozygous mice have normal T cell development). Noticeable within the WES of this mouse in addition to the RAG1 mutation were the other 1385 *de novo* mutations (compared to the mouse reference genome mm10), 95 of which were homozygous.

Lesson 3: Notice results that do not make sense. Know when an unexpected result may be interesting. It was unexpected that Prf1$^{-/-}$IFNγ$^{-/-}$ would have a T cell developmental block. Yet, the most unexpected part of the discovery of the RAG1 mutation, was that T and B cells existed in this mouse at all. It would have been very easy to use this result to develop a genotyping strategy to screen out the mutation from our Prf1$^{-/-}$IFNγ$^{-/-}$ mice, but instead we were able to develop this project into what we hope will be a significant contribution to the RAG and T cell development fields.

Too often in publication it seems that 'spin' hides the true pathways to discovery of novel ideas. In fact, in an early telling of this exact story, in this chronological manner, I was told "this is interesting, but you cannot tell the story this way". The need to have a perfectly hypothesis-driven discovery can mask some of the most interesting parts of many discovery stories, and I think it important that science does not hide the ability to discover by chance, à la Alexander Fleming, for fear that this diminishes the achievement.

4.7: Figures

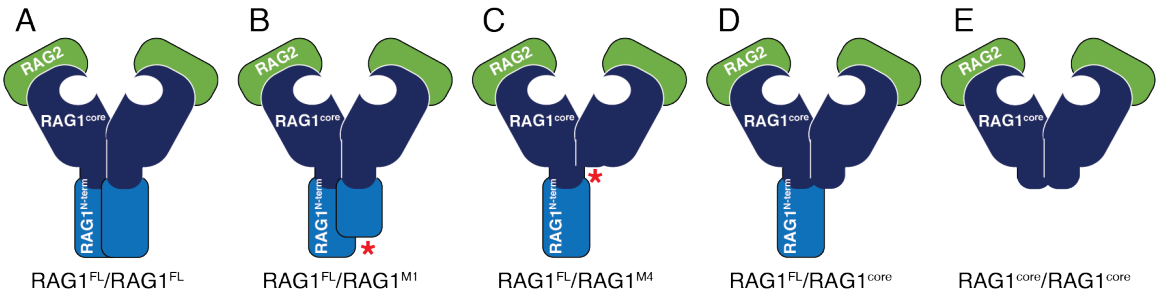


Figure 4-1: Cartoons depicting alternative RAG1 isoforms in hetero-tetrameric RAG complex

Potential for different stoichiometries of RAG1 isoforms to form part of the $RAG1_2RAG2_2$ heterotetrameric complex. Not representative of all possible combinations, but some potential examples. (A) RAG1 full-length (FL) only. (B) RAG1-FL with RAG1 translated from M1. (C) RAG1-FL with RAG1 translated from M4. (D) RAG1-FL co-expressed with RAG1-core. (E) RAG1-core only.

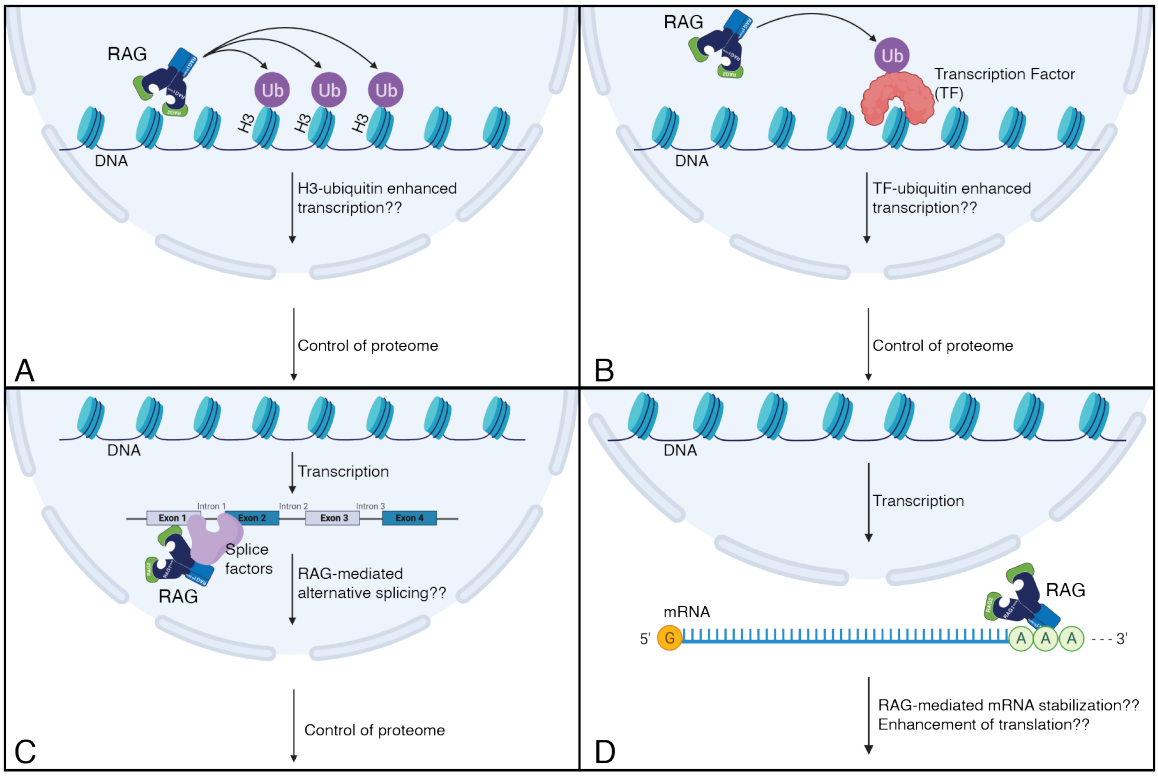


Figure 4-2: Speculative mechanisms of RAG1 N-terminal domain mediated control of the proteome

(A) RAG1 endogenous E3 ligase activity, or in association with VprBP, could lead to ubiquitylation of histones, such as H3, to enhance or repress transcriptional programs to modulate the proteome. (B) RAG1 endogenous E3 ligase activity, or in association with VprBP, could lead to ubiquitylation of transcription factors to enhance or repress transcriptional programs to modulate the proteome. (C) The RAG1 N-terminal domain binds an RNA-splicing factor (57). It is possible that alternative splicing is modulated by RAG1 N-terminal domain interactions with splice factors. (D) The most unlikely possibility is that RAG1 acts to stabilize mRNA or enhance translation in the cytoplasm via its RNA-binding ability.

APPENDIX

Materials and Methods

Mice

All mice within this study were housed, bred, and used under pathogen-free conditions at the Children's Hospital of Philadelphia (CHOP). Both female and male mice were used for all experiments. Experimental mice were euthanized by CO_2 exposure followed by cervical dislocation. Animal husbandry and experiments were performed in accordance with national guidelines and approved by the CHOP Institutional Animal Care and Use Committee. $RAG1^{WT}$ (Purchased from Jackson Laboratories), $RAG1^{NX}$ (generated in our facility), $RAG1^{PG}$ (a gift from Dr. Jessica Jones (Georgetown University) via Dr. David Schatz (Yale)), $RAG1^{-/-}$ (Purchased from Jackson Laboratories), and OT1 (Purchased from Jackson Laboratories) mice are all on a C57BL/6 background. $RAG1^{DA}$ (178), $Artemis^{-/-}Bcl2^{tg}$ (179), and VβNT (148) mice were all on a mixed 129S1/SvImJ and C57BL/6 background, and crossed to $RAG1^{WT}$, $RAG1^{core}$, or $RAG1^{PG}$ mice in the Bassing Lab.

Whole Exome analysis and Sanger sequencing

Genomic DNA was isolated from $RAG1^{NX}$ spleen using the DNeasy Blood and Tissue Kit (Qiagen) per manufacturer's instructions. Exon capture for whole exome sequencing was performed using the Agilent SureSelect XT Mouse All Exon kit. Sequencing was performed on the Illumina HiSeq 4000 to produce 150-bp paired-end reads with an average depth of 100X. Sequence reads were aligned to the reference mouse genome (GRCm38/mm10) using Novoalign (V3.03.01; http://www.novocraft.com). Picard was

used for marking duplicates then variants were called using GATK's HaplotypeCaller. Single nucleotide variants and insertions/deletions were functionally annotated with SnpEff (http://snpeff.sourceforge.net) and filtered to retain only moderate and high effect variants. After applying quality filters and excluding variants that have been reported in the Single Nucleotide Polymorphism Database, 299 variants remained. Genes were then annotated using gene ontologies to identify variants in genes involved in B and T cell development. A homozygous stop gain mutation was identified in Rag1 (NM_009019.2; p.Gln60*) that met all filtering and quality control criteria.

Flow cytometric analysis

Single cell suspensions of all organs were ACK lysed to remove erythrocytes before they were stained with LIVE/DEAD fixable viability dye (Life Technologies) and antibodies against respective surface antigens (BD Pharmingen, eBioscience, and Biolegend). For intracellular HA, FLAG, GFP, and FoxP3 staining, cells were stained using the eBioscience FoxP3 kit according to manufacturer's instructions. CFSE labeling of cells was performed as per manufacturer's instructions (Invitrogen). All samples were acquired on a MACSQuant flow cytometer (Miltenyi Biotec) or LSRII Fortessa (BD) and analyzed using Flowjo software version 10.5.3 (Tree Star).

Intracellular staining

Staining of intracellular cytokines: Cells (10^6) were cultured in the absence or presence of 50ng/ml PMA (Sigma), 1μg/ml Ionomycin (Cell Signaling Technology), with 2μg/ml brefeldin A (Sigma) and 2μM monensin (eBioscience) for 4 hours at 37°C. After staining

for LIVE/DEAD and for surface antigens as described earlier, cells were stained for the respective cytokines using the Cytofix/Cytoperm kit (BD Bioscience).

Staining of transcription factors: Cells were stained for respective transcription factors using FoxP3/Transcription Factor Staining kit (eBioscience/Thermofisher)

qPCR of VDJ rearrangements in DN3 and DP thymocytes

DN3 cells for TCRβ rearrangements: Thymocytes were isolated, stained with PE-labelled CD4, CD8, CD11b, CD11c, NK1.1, Gr1, and Ter119. Non-labelled cells were enriched by MACS depletion using anti-PE microbeads and LS columns (Miltenyi). Enriched cells were then stained with CD4, CD8, CD44, and CD25. DN3 cells were sorted using a FACSAria Fusion. Genomic DNA was extracted from the sorted DN3 cells using the DNeasy Blood and Tissue kit (QIAGEN). A quantitative PCR assay to measure Vβ-Dβ1-Jβ1 and Vβ-Dβ2-Jβ2 rearrangement frequencies was designed with a panel of primers specific for each functional Vβ paired with a probe (FAM, HEX) specific for either Jβ1.1 or Jβ2.1, respectively. Rearrangements were measured by TaqMan qPCR with PCR conditions according to the manufacturer's instructions (PrimeTime IDTDNA). PCR analysis of TRDV1-JD1 and CD19 were used for normalization. Primers and probes described previously (220).

DP cells for TCRα rearrangements: Thymocytes were stained for CD4 and CD8 and live DP cells were sorted. Genomic DNA was isolated as above. Rearrangements were determined via qPCR using conditions and primers previously described, using β2m as a housekeeping gene (221, 222).

Antinuclear antibodies

Hep2 ANA slides (MBL International) were incubated with mouse serum diluted 100fold in PBS for 15minutes in the dark. Slides were washed with PBS and stained with DAPI (Thermofisher), and anti-mouse Ig-AF488 (Jackson Immunoresearch Laboratories). Images were acquired on a at 20x Plan Apo 1.4 NA objective on a spinning disk confocal system (UltraView ERS 6; PerkinElmer, Waltham, MA) equipped with an ORCA-ER camera (Hamamatsu Photonics, Bridgewater, NJ) and velocity software (v6.1.1; PerkinElmer). Instrument settings were fixed for all images, and researchers were blinded to the sample identification. Nuclear AF488 intensity was quantified using FIJI software, using DAPI to designate nuclear region of interest. A minimum of 30 nuclei were analyzed per well and the average AF488 intensity was reported.

Cloning and RAG1 mutagenesis

FLAG-RAG1-HA was amplified by PCR using RAG1-HA template from the Bassing lab, and cloned into the NotI and XhoI sites of pZHK:CMV-IRES-GFP (223) to create the base pZHK:CMV-FLAG-RAG1(WT)-HA-IRES-GFP construct. For mutagenesis of methionines and Q60X, RAG1 was subcloned into puc18 and mutagenesis was performed using Quickchange Lightning Kit (Agilent). Correctly mutated fragments were swapped back in to the pZHK vector for expression.

Transfection of 293T cells

293T cells were grown in 6 well plates to 70-80% confluency. 2μg of the respective RAG1 constructs were transfected in using TransIT-Lenti reagent (MIRUS) according to

manufacturer's instructions. Cells were harvested 16-20hours post transfection for flow cytometry or western blot.

SDS-PAGE and Western Blotting

Native RAG1 western blot was performed as previously described (224) using RAG1 antibody from David Schatz (Yale). For RAG1 expressed in 293T cells, transfected cells were lysed using M-PER lysis buffer (ThermoFisher) with protease inhibitor cocktail (ThermoFisher). 15ug of protein was boiled in loading buffer with 5% β-ME and run on a 4-12% SDS-PAGE gel before transfer to nitrocellulose membrane (Bio-Rad). Membranes were probed with anti-HA (Clone 3F10 – Sigma) and anti-FLAG (Clone M2 – Sigma), then AF700 and AF780 conjugated secondary antibodies, before imaging on a Licor Odyssey.

Co-immunoprecipitation of RAG1/RAG2

293T cells were co-transfected with RAG1 and RAG2 vectors as described in the text. To immunoprecipitate, the GFP-trap agarose kit from Chromotek was used and procedure was as the manufacturer described. Precipitated protein was run on an SDS-PAGE gel and Western blotting was performed as above.

RT-qPCR for mRNA expression

Respective cell populations (described in text) were sorted into RLT buffer containing 2-mercaptoethanol using a FACSaria Fusion cell sorter. RNA was isolated using Qiagen RNeasy Mini kit according to manufacturer's instructions. Isolated RNA was converted to cDNA with Superscript III First Strand Synthesis kit utilizing random hexamers according to manufacturer's instructions (Invitrogen). cDNA was subjected to qPCR using Power

SYBR Green kit (Applied Biosystems) and HPRT, ZAP70, and SYK primers (Qiagen). Relative expression was calculated using the ddCt method, using HPRT as a housekeeping gene, and relevant calibrator sample (explained in figure legends for respective samples).

ZAP70 overexpression

GP+ E-86 ecotropic packaging cell lines that stably produce MLV-based ZAP70-IRES-GFP ($ZAP70^{OE}$), or GFP-only control retrovirus were a kind gift from Dr. Naomi Taylor (National Cancer Institute, NIH) (225). Cells were cultured in DMEM + 10% FBS and virus was harvested from cells grown to near confluence. Virus was concentrated using Retro-X concentrator (Takara) following manufacturers guidelines and frozen at -80^{o}C until use.

Bone marrow cells were stained with PE-conjugated lineage antibodies (CD5, B220, NK1.1, CD11b, Gr1, Ly6G, Ter119) and then incubated with magnetic anti-PE beads (Miltenyi). Cells were then run through LS columns (Miltenyi) to enrich for $Lineage^{-}$ cells. Cells were then stained with Sca-1 and cKIT, and $Lineage^{-}Sca1^{+}cKIT^{int/+}$ (LSK) cells were sorted using a FACSaria Fusion.

LSK cells were grown in RPMI + 10% FBS + 1% PSG with 50ng/ml rIL-6, 20ng/ml rIL-3, 50ng/ml rSCF (Biolegend) for 4 days. On day 2 and day 3 cells were spin-transduced with $ZAP70^{OE}$ or control retrovirus with 8μg/ml polybrene (800g x 2hours x 32^{o}C). On day 4 cells were transferred into lethally irradiated host mice (950Rads – Xrad irradiator). 0.5E6 $RAG1^{-/-}$ cells were also transferred to each mouse to ensure survival.

OT1 : TAP1$^{-/-}$ Chimeras

Bone marrow from RAG1^{WT}OT1^{+}CD45.1-2 and RAG1^{PG}OT1^{+}CD45.2 mice was isolated and combined in a 1:1 ratio. Cells were then transferred to lethally irradiated (950Rad – Xrad irradiator) host mice and mice were allowed to reconstitute for 4-8 weeks prior to experiment.

LCMV infections

Mice were infected intraperitoneally with 2x10^{5} plaque-forming units of LCMV-Armstrong and euthanized at indicated timepoints. Virus and gp33-tetramer were kindly provided by the Wherry Lab (University of Pennsylvania).

Transfer model of colitis

This model was performed as previously described (154). Briefly, naïve CD4+ T cells were sorted from respective donor mice (CD90.2+, CD4+, CD62L+, CD44+, CD25-) on a FACSaria Fusion cell sorter. 0.5E6 cells were transferred in 100ul into each of the donor mice via intraperitoneal injection. Mice were weighed weekly for the first 4 weeks, and then twice weekly for the next 4 weeks. Mice were euthanized 8 weeks post-transfer, unless severe morbidity enforced premature euthanasia (as dictated by CHOP animal facility veterinarians), or natural death occurred.

ELISAs

Blood was isolated by terminal cardiac puncture and allowed to clot at 4^{o}C. Cells were centrifuged at 15000g for 15 minutes and serum isolated and frozen at -20^{o}C. IL-6 and IFNγ ELISAs were performed according to manufacturer's instructions (BD Biosciences).

In vitro T cell differentiation

Splenocytes from RAG1$^{-/-}$ hosts were irradiated with 2500Rads (X-rad irradiator) and 3E5 cells were seeded into each well of a 96 well plate. Naïve T cells (CD90.2+, CD4+, CD62L+, CD44+, CD25-) were sorted from respective mice on a FACSaria Fusion cell sorter, and 3E4 sorted CD4s were added to each well containing irradiated feeder cells. All cells were cultured with 10μg/ml anti-CD3 (Biolegend) and 3μg/ml anti-CD28 (Biolegend) in RPMI + 10% FBS + 1% PSG + 1x NEAA (Invitrogen), 1x Sodium Pyruvate (Invitrogen), and 0.001% 2-mercaptoethanol. The following cytokines/blocking antibodies and concentrations were used to skew towards respective Th-subsets:

Culturing Condition	Reagent	Supplier	Concentration
Th0	rhIL2	Peprotech	30U/ml
Th1	rmIL12	Peprotech	10ng/ml
	anti-mIL4	Biolegend	10μg/ml
	rhIL2	Peprotech	30U/ml
Th2	rmIL4	Biolegend	10ng/ml
	anti-mIFNγ	Biolegend	10μg/ml
	anti-mIL12	Biolegend	10μg/ml
	rhIL2	Peprotech	30U/ml
Th17	rmIL6	Peprotech	20ng/ml
	rmTGFb1	Biolegend	1ng/ml
	anti-mIL4	Biolegend	10μg/ml
	anti-mIL12	Biolegend	10μg/ml
	rhIL2	Peprotech	30U/ml
Treg	rmTGFb1	Biolegend	1ng/ml
	anti-mIL4	Biolegend	10μg/ml
	anti-mIL12	Biolegend	10μg/ml
	rhIL2	Peprotech	30U/ml

Statistical analysis

All data was analyzed in Graphpad Prism 8 using statistical tests indicated in the figure legends. Error bars indicate mean +/- SEM unless otherwise stated. n.s.=$p>0.05$, *$p<0.05$, **$p<0.01$, ***$p<0.0001$, ****$p<0.0001$.

BIBLIOGRAPHY

1. Bergh O, Borsheim KY, Bratbak G, Heldal M. High abundance of viruses found in aquatic environments. Nature. 1989;340(6233):467-8.

2. Suttle CA. Viruses in the sea. Nature. 2005;437(7057):356-61.

3. Weinbauer MG. Ecology of prokaryotic viruses. FEMS Microbiol Rev. 2004;28(2):127-81.

4. Microbiology by numbers. Nat Rev Microbiol. 2011;9(9):628.

5. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337(6096):816-21.

6. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339(6121):819-23.

7. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol. 2005;60(2):174-82.

8. Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology. 2005;151(Pt 3):653-63.

9. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 2005;151(Pt 8):2551-61.

10. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008;321(5891):960-4.

11. Kaufmann SHE. Remembering Emil von Behring: from Tetanus Treatment to Antibody Cooperation with Phagocytes. mBio. 2017;8(1).

12. Behring E, Kitasato S. Ueber das Zustandekommen der Diphtherie-Immunität und der Tetanus-Immunität bei Thieren. Dtsch Med Wochenschrift. 1890;49:1113–4.

13. Ehrlich P. Die Seitenkettentheorie und ihre Gegner. Münchner Med Wochenschr. 1901:2123-4.

14. Silverstein AM. Paul Ehrlich's passion: the origins of his receptor immunology. Cell Immunol. 1999;194(2):213-21.

15. Jerne NK. THE NATURAL-SELECTION THEORY OF ANTIBODY FORMATION. PNAS. 1955;41(11):849-57.

16. Burnet FM. A Modification of Jerne's Theory of Antibody Production using the Concept of Clonal Selection. Australian Journal of Science. 1957;20(3):67-9.

17. J.C. V, al e. The sequence of the human genome. Science. 2001;291(5507):1304-51.

18. Hedrick SM, Germain RN, Bevan MJ, Dorf M, Engel I, Fink P, et al. Rearrangement and transcription of a T-cell receptor beta-chain gene in different T-cell subsets. PNAS. 1985;82(2):531-5.

19. Jenkins MK, Chu HH, McLachlan JB, Moon JJ. On the composition of the preimmune repertoire of T cells specific for Peptide-major histocompatibility complex ligands. Annu Rev Immunol. 2010;28:275-94.

20. Arstila TP, Casrouge A, Baron V, Even J, Kanellopoulos J, Kourilsky P. A direct estimate of the human alphabeta T cell receptor diversity. Science. 1999;286(5441):958-61.

21. Qi Q, Liu Y, Cheng Y, Glanville J, Zhang D, Lee JY, et al. Diversity and clonal selection in the human T-cell repertoire. Proc Natl Acad Sci U S A. 2014;111(36):13139-44.

22. Briney B, Inderbitzin A, Joyce C, Burton DR. Commonality despite exceptional diversity in the baseline human antibody repertoire. Nature. 2019;566(7744):393-7.

23. Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. Nature. 1988;334(6181):395-402.

24. Hozumi N, Tonegawa S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. Proc Natl Acad Sci U S A. 1976;73(10):3628-32.

25. Brack C, Hirama M, Lenhard-Schuller R, Tonegawa S. A complete immunoglobulin gene is created by somatic recombination. Cell. 1978;15(1):1-14.

26. Tonegawa S. Somatic generation of antibody diversity. Nature. 1983;302(5909):575-81.

27. Rabbitts TH, Lefranc MP, Stinson MA, Sims JE, Schroder J, Steinmetz M, et al. The chromosomal location of T-cell receptor genes and a T cell rearranging gene: possible correlation with specific translocations in human T cell leukaemia. Embo j. 1985;4(6):1461-5.

28. Isobe M, Erikson J, Emanuel BS, Nowell PC, Croce CM. Location of gene for beta subunit of human T-cell receptor at band 7q35, a region prone to rearrangements in T cells. Science. 1985;228(4699):580-2.

29. Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. Nature Reviews Immunology. 2011;11(4):251-63.

30. Kallenbach S, Doyen N, Fanton d'Andon M, Rougeon F. Three lymphoid-specific factors account for all junctional diversity characteristic of somatic assembly of T-cell receptor and immunoglobulin genes. Proc Natl Acad Sci U S A. 1992;89(7):2799-803.

31. Roth DB. V(D)J Recombination: Mechanism, Errors, and Fidelity. Microbiol Spectr. 2014;2(6).

32. Jones JM, Gellert M. Ordered assembly of the V(D)J synaptic complex ensures accurate recombination. Embo j. 2002;21(15):4162-71.

33. Mundy CL, Patenge N, Matthews AGW, Oettinger MA. Assembly of the RAG1/RAG2 Synaptic Complex. Mol Cell Biol. 222002. p. 69-77.

34. Bassing CH, Alt FW, Hughes MM, D'Auteuil M, Wehrly TD, Woodman BB, et al. Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. Nature. 2000;405(6786):583-6.

35. van-Gent DC, Ramsden DA, Gellert M. The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. Cell. 1996;85(1):107-13.

36. Eastman QM, Leu TM, Schatz DG. Initiation of V(D)J recombination in vitro obeying the 12/23 rule. Nature. 1996;380(6569):85-8.

37. Krangel MS. Mechanics of T cell receptor gene rearrangement. Curr Opin Immunol. 2009;21(2):133-9.

38. van-Gent DC, Hiom K, Paull TT, Gellert M. Stimulation of V(D)J cleavage by high mobility group proteins. EMBO J. 1997;16(10):2665-70.

39. Schatz DG, Swanson PC. V(D)J recombination: mechanisms of initiation. Annu Rev Genet. 2011;45:167-202.

40. Ma Y, Pannicke U, Schwarz K, Lieber MR. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell. 2002;108(6):781-94.

41. Lieber MR. The mechanism of human nonhomologous DNA end joining. J Biol Chem. 2008;283(1):1-5.

42. Lu H, Schwarz K, Lieber MR. Extent to which hairpin opening by the Artemis:DNA-PKcs complex can contribute to junctional diversity in V(D)J recombination. Nucleic Acids Res. 2007;35(20):6917-23.

43. Komori T, Okada A, Stewart V, Alt FW. Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. Science. 1993;261(5125):1171-5.

44. Notarangelo LD, Kim MS, Walter JE, Lee YN. Human RAG mutations: biochemistry and clinical implications. Nat Rev Immunol. 2016;16(4):234-46.

45. Mombaerts P, Iacomini J, Johnson RS, Herrup K, Tonegawa S, Papaioannou VE. RAG-1-deficient mice have no mature B and T lymphocytes. Cell. 1992;68(5):869-77.

46. Zhang Y, Cheng TC, Huang G, Lu Q, Surleac MD, Mandell JD, et al. Transposon molecular domestication and the evolution of the RAG recombinase. Nature. 2019;569(7754):79-84.

47. Dudley DD, Sekiguchi J, Zhu C, Sadofsky MJ, Whitlow S, DeVido J, et al. Impaired V(D)J recombination and lymphocyte development in core RAG1-expressing mice. The Journal of experimental medicine. 2003;198(9):1439-50.

48. Horowitz JE, Bassing CH. Non-Core RAG1 Regions Promote Vβ Rearrangements and αβ T Cell Development by Overcoming Inherent Inefficiency of Vβ Recombination Signal Sequences. J Immunol. 2014;192(4):1609-19.

49. Jones JM, Bhattacharyya A, Simkus C, Vallieres B, Veenstra TD, Zhou M. The RAG1 V(D)J recombinase/ubiquitin ligase promotes ubiquitylation of acetylated, phosphorylated histone 3.3. Immunol Lett. 2011;136(2):156-62.

50. Jones JM, Gellert M. Autoubiquitylation of the V(D)J recombinase protein RAG1. Proc Natl Acad Sci U S A. 2003;100(26):15446-51.

51. Yurchenko V, Xue Z, Sadofsky M. The RAG1 N-terminal domain is an E3 ubiquitin ligase. Genes Dev. 2003;17(5):581-5.

52. Grazini U, Zanardi F, Citterio E, Casola S, Goding CR, McBlane F. The RING domain of RAG1 ubiquitylates histone H3: a novel activity in chromatin-mediated regulation of V(D)J joining. Mol Cell. 2010;37(2):282-93.

53. Simkus C, Anand P, Bhattacharyya A, Jones JM. Biochemical and folding defects in a RAG1 variant associated with Omenn syndrome. J Immunol. 2007;179(12):8332-40.

54. Kassmeier MD, Mondal K, Palmer VL, Raval P, Kumar S, Perry GA, et al. VprBP binds full-length RAG1 and is required for B-cell development and V(D)J recombination fidelity. EMBO J. 2012;31(4):945-58.

55. Coster G, Gold A, Chen D, Schatz DG, Goldberg M. A dual interaction between the DNA damage response protein MDC1 and the RAG1 subunit of the V(D)J recombinase. J Biol Chem. 2012;287(43):36488-98.

56. Raval P, Kriatchko AN, Kumar S, Swanson PC. Evidence for Ku70/Ku80 association with full-length RAG1. Nucleic Acids Res. 2008;36(6):2060-72.

57. Maitra R, Sadofsky MJ. A WW-like module in the RAG1 N-terminal domain contributes to previously unidentified protein-protein interactions. Nucleic Acids Res. 2009;37(10):3301-9.

58. Miller JF. Immunological function of the thymus. Lancet. 1961;2(7205):748-9.

59. McDonald-McGinn DM, Sullivan KE, Marino B, Philip N, Swillen A, Vorstman JA, et al. 22q11.2 deletion syndrome. Nat Rev Dis Primers. 2015;1:15071.

60. Laurenti E, Göttgens B. From haematopoietic stem cells to complex differentiation landscapes. Nature. 2018;553(7689):418-26.

61. Bhandoola A, von Boehmer H, Petrie HT, Zuniga-Pflucker JC. Commitment and developmental potential of extrathymic and intrathymic T cell precursors: plenty to choose from. Immunity. 2007;26(6):678-89.

62. Aifantis I, Buer J, von Boehmer H, Azogui O. Essential role of the pre-T cell receptor in allelic exclusion of the T cell receptor beta locus. Immunity. 1997;7(5):601-7.

63. Brady BL, Steinel NC, Bassing CH. Antigen Receptor Allelic Exclusion: An Update and Reappraisal. J Immunol. 2010;185(7):3801-8.

64. Carleton M, Ruetsch NR, Berger MA, Rhodes M, Kaptik S, Wiest DL. Signals transduced by CD3epsilon, but not by surface pre-TCR complexes, are able to induce maturation of an early thymic lymphoma in vitro. J Immunol. 1999;163(5):2576-85.

65. Irving BA, Alt FW, Killeen N. Thymocyte development in the absence of pre-T cell receptor extracellular immunoglobulin domains. Science. 1998;280(5365):905-8.

66. Hoffman ES, Passoni L, Crompton T, Leu TM, Schatz DG, Koff A, et al. Productive T-cell receptor beta-chain gene rearrangement: coincident regulation of cell cycle and clonality during development in vivo. Genes Dev. 1996;10(8):948-62.

67. Klein L, Kyewski B, Allen PM, Hogquist KA. Positive and negative selection of the T cell repertoire: what thymocytes see and don't see. Nat Rev Immunol. 2014;14(6):377-91.

68. Krueger A, Zietara N, Lyszkiewicz M. T Cell Development by the Numbers. Trends Immunol. 2017;38(2):128-39.

69. Kincaid EZ, Murata S, Tanaka K, Rock KL. Specialized proteasome subunits play an essential role in thymic selection of CD8+ T cells. Nat Immunol. 2016;17(8):938-45.

70. Kurobe H, Liu C, Ueno T, Saito F, Ohigashi I, Seach N, et al. CCR7-dependent cortex-to-medulla migration of positively selected thymocytes is essential for establishing central tolerance. Immunity. 2006;24(2):165-77.

71. Forster R, Davalos-Misslitz AC, Rot A. CCR7 and its ligands: balancing immunity and tolerance. Nat Rev Immunol. 2008;8(5):362-71.

72. Derbinski J, Schulte A, Kyewski B, Klein L. Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. Nat Immunol. 2001;2(11):1032-9.

73. Abramson J, Giraud M, Benoist C, Mathis D. Aire's partners in the molecular control of immunological tolerance. Cell. 2010;140(1):123-35.

74. Weinreich MA, Hogquist KA. Thymic emigration: when and how T cells leave home. J Immunol. 2008;181(4):2265-70.

75. Anderson MS, Venanzi ES, Klein L, Chen Z, Berzins SP, Turley SJ, et al. Projection of an immunological self shadow within the thymus by the aire protein. Science. 2002;298(5597):1395-401.

76. Heino M, Peterson P, Kudoh J, Shimizu N, Antonarakis SE, Scott HS, et al. APECED mutations in the autoimmune regulator (AIRE) gene. Hum Mutat. 2001;18(3):205-11.

77. Sakaguchi N, Takahashi T, Hata H, Nomura T, Tagami T, Yamazaki S, et al. Altered thymic T-cell selection due to a mutation of the ZAP-70 gene causes autoimmune arthritis in mice. Nature. 2003;426(6965):454-60.

78. Jordan MS, Smith JE, Burns JC, Austin JET, Nichols KE, Aschenbrenner AC, et al. Complementation in trans of altered thymocyte development in knock-in mice expressing mutant forms of SLP76. Immunity. 2008;28(3):359-69.

79. Hsu LY, Tan YX, Xiao Z, Malissen M, Weiss A. A hypomorphic allele of ZAP-70 reveals a distinct thymic threshold for autoimmune disease versus autoimmune reactivity. J Exp Med. 2009;206(11):2527-41.

80. Davey GM, Schober SL, Endrizzi BT, Dutcher AK, Jameson SC, Hogquist KA. Preselection thymocytes are more sensitive to T cell receptor stimulation than mature T cells. J Exp Med. 1998;188(10):1867-74.

81. Lucas B, Stefanova I, Yasutomo K, Dautigny N, Germain RN. Divergent changes in the sensitivity of maturing T cells to structurally related ligands underlies formation of a useful T cell repertoire. Immunity. 1999;10(3):367-76.

82. Stritesky GL, Jameson SC, Hogquist KA. Selection of self-reactive T cells in the thymus. Annu Rev Immunol. 2012;30:95-114.

83. Gascoigne NR, Rybakin V, Acuto O, Brzostek J. TCR Signal Strength and T Cell Development. Annu Rev Cell Dev Biol. 2016;32:327-48.

84. Hogquist KA, Jameson SC. The self-obsession of T cells: how TCR signaling thresholds affect fate 'decisions' and effector function. Nat Immunol. 2014;15(9):815-23.

85. Sant'Angelo DB, Janeway CA, Jr. Negative selection of thymocytes expressing the D10 TCR. Proc Natl Acad Sci U S A. 2002;99(10):6931-6.

86. Pircher H, Rohrer UH, Moskophidis D, Zinkernagel RM, Hengartner H. Lower receptor avidity required for thymic clonal deletion than for effector T-cell function. Nature. 1991;351(6326):482-5.

87. Fu G, Rybakin V, Brzostek J, Paster W, Acuto O, Gascoigne NR. Fine-tuning T cell receptor signaling to control T cell development. Trends Immunol. 2014;35(7):311-8.

88. Shiloh Y. ATM and related protein kinases: safeguarding genome integrity. Nat Rev Cancer. 2003;3(3):155-68.

89. Rouse J, Jackson SP. Interfaces between the detection, signaling, and repair of DNA damage. Science. 2002;297(5581):547-51.

90. Innes CL, Heinloth AN, Flores KG, Sieber SO, Deming PB, Bushel PR, et al. ATM requirement in gene expression responses to ionizing radiation in human lymphoblasts and fibroblasts. Mol Cancer Res. 2006;4(3):197-207.

91. Meek DW. The p53 response to DNA damage. DNA Repair (Amst). 2004;3(8-9):1049-56.

92. Bredemeyer AL, Helmink BA, Innes CL, Calderon B, McGinnis LM, Mahowald GK, et al. DNA double strand breaks activate a multi-functional genetic program in developing lymphocytes. Nature. 2008;456(7223):819-23.

93. Bednarski JJ, Pandey R, Schulte E, White LS, Chen BR, Sandoval GJ, et al. RAG-mediated DNA double-strand breaks activate a cell type-specific checkpoint to inhibit pre-B cell receptor signals. J Exp Med. 2016;213(2):209-23.

94. Bednarski JJ, Sleckman BP. At the intersection of DNA damage and immune responses. Nat Rev Immunol. 2019;19(4):231-42.

95. Fox CJ, Hammerman PS, Cinalli RM, Master SR, Chodosh LA, Thompson CB. The serine/threonine kinase Pim-2 is a transcriptionally regulated apoptotic inhibitor. Genes Dev. 2003;17(15):1841-54.

96. Villa A, Santagata S, Bozzi F, Imberti L, Notarangelo LD. Omenn syndrome: a disorder of Rag1 and Rag2 genes. J Clin Immunol. 1999;19(2):87-97.

97. Villa A, Santagata S, Bozzi F, Giliani S, Frattini A, Imberti L, et al. Partial V(D)J recombination activity leads to Omenn syndrome. Cell. 1998;93(5):885-96.

98. Villa A, Marrella V, Rucci F, Notarangelo L. Genetically determined lymphopenia and autoimmune manifestations. Current Opinion in Immunology. 2008;20(3):318-24.

99. Santagata S, Gomez C, Sobacchi C, Bozzi F, Abinum M, Pasic S, et al. N-terminal RAG1 frameshift mutations in Omenn's syndrome: internal methionine usage leads to partial V(D)J recombination activity and reveals a fundamental role in vivo for the N-terminal domains. PNAS. 2000;97(26):14572-7.

100. de Klerk E, t Hoen PA. Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. Trends Genet. 2015;31(3):128-39.

101. Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell. 2011;147(4):789-802.

102. Ingolia NT. Ribosome Footprint Profiling of Translation throughout the Genome. Cell. 2016;165(1):22-33.

103. Lee S, Liu B, Lee S, Huang S-X, Shen B, Qian S-B. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. PNAS. 2012.

104. Spriggs KA, Bushell M, Willis AE. Translational regulation of gene expression during conditions of cell stress. Mol Cell. 2010;40(2):228-37.

105. Hellen CU, Sarnow P. Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev. 2001;15(13):1593-612.

106. Candeias MM, Powell DJ, Roubalova E, Apcher S, Bourougaa K, Vojtesek B, et al. Expression of p53 and p53/47 are controlled by alternative mechanisms of messenger RNA translation initiation. Oncogene. 2006;25(52):6936-47.

107. Ray PS, Grover R, Das S. Two internal ribosome entry sites mediate the translation of p53 isoforms. EMBO Rep. 2006;7(4):404-10.

108. Kochetov AV. Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. Bioessays. 2008;30(7):683-91.

109. Salat-Canela C, Sese M, Peula C, Ramon y Cajal S, Aasen T. Internal translation of the connexin 43 transcript. Cell Commun Signal. 2014;12:31.

110. Kozak M. New ways of initiating translation in eukaryotes? Mol Cell Biol. 2001;21(6):1899-907.

111. Smyth JW, Shaw RM. Autoregulation of connexin43 gap junction formation by internally translated isoforms. Cell Rep. 2013;5(3):611-8.

112. Zydowicz-Machtel P, Swiatkowska A, Popenda L, Gorska A, Ciesiolka J. Variants of the 5'-terminal region of p53 mRNA influence the ribosomal scanning and translation efficiency. Sci Rep. 2018;8(1):1533.

113. van der Lugt NM, Domen J, Verhoeven E, Linders K, van der Gulden H, Allen J, et al. Proviral tagging in E mu-myc transgenic mice lacking the Pim-1 proto-oncogene leads to compensatory activation of Pim-2. EMBO J. 1995;14(11):2536-44.

114. Rubtsova MP, Sizova DV, Dmitriev SE, Ivanov DS, Prassolov VS, Shatsky IN. Distinctive properties of the 5'-untranslated region of human hsp70 mRNA. J Biol Chem. 2003;278(25):22350-6.

115. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. Cell. 1989;59(6):1035-48.

116. Carpenter AC, Bosselut R. Decision checkpoints in the thymus. Nat Immunol. 2010;11(8):666-73.

117. Khiong K, Murakami M, Kitabayashi C, Ueda N, Sawa S, Sakamoto A, et al. Homeostatically proliferating CD4 T cells are involved in the pathogenesis of an Omenn syndrome murine model. J Clin Invest. 2007;117(5):1270-81.

118. Marrella V, Poliani PL, Casati A, Rucci F, Frascoli L, Gougeon ML, et al. A hypomorphic R229Q Rag2 mouse mutant recapitulates human Omenn syndrome. J Clin Invest. 1172007. p. 1260-9.

119. Watanabe H, Numata K, Ito T, Takagi K, Matsukawa A. Innate immune response in Th1- and Th2-dominant mouse strains. Shock. 2004;22(5):460-6.

120. Rubtsova K, Rubtsov AV, Cancro MP, Marrack P. Age-associated B cells: A T-bet dependent effector with roles in protective and pathogenic immunity. J Immunol. 2015;195(5):1933-7.

121. McMahan CJ, Difilippantonio MJ, Rao N, Spanopoulou E, Schatz DG. A basic motif in the N-terminal region of RAG1 enhances V(D)J recombination activity. Mol Cell Biol. 1997;17(8):4544-52.

122. Noderer W, Flockhart R, Bhaduri A, Diaz de Arce A, Zhang J, Khavari P, et al. Quantitative analysis of mammalian translation initiation sites by FACS-seq. Mol Syst Biol. 2014;10(748).

123. Goldrath AW, Luckey CJ, Park R, Benoist C, Mathis D. The molecular program induced in T cells undergoing homeostatic proliferation. PNAS. 2004;101(48):16885-90.

124. Hsieh CS, Liang Y, Tyznik AJ, Self SG, Liggitt D, Rudensky AY. Recognition of the peripheral self by naturally arising CD25+ CD4+ T cell receptors. Immunity. 2004;21(2):267-77.

125. Curotto de Lafaille M, Lino A, Kutchukhidze N, Lafaille J. CD25- T cells generate CD25+Foxp3+ regulatory T cells by peripheral expansion. J Immunol. 2004;173(12):7259-68.

126. Jiang W, Anderson MS, Bronson R, Mathis D, Benoist C. Modifier loci condition autoimmunity provoked by Aire deficiency. J Exp Med. 2005;202(6):805-15.

127. Jin Y, Lee A, Oh JH, Lee HW, Ha SJ. The R229Q mutation of Rag2 does not characterize severe immunodeficiency in mice. Sci Rep. 2019;9(1):4415.

128. Liang HE, Hsu LY, Cado D, Schlissel MS. Variegated transcriptional activation of the immunoglobulin kappa locus in pre-b cells contributes to the allelic exclusion of light-chain expression. Cell. 2004;118(1):19-29.

129. Perry RP, Kelley DE, Coleclough C, Seidman JG, Leder P, Tonegawa S, et al. Transcription of mouse kappa chain genes: implications for allelic exclusion. Proc Natl Acad Sci U S A. 1980;77(4):1937-41.

130. Schatz DG, Baltimore D. Stable expression of immunoglobulin gene V(D)J recombinase activity by gene transfer into 3T3 fibroblasts. Cell. 1988;53(1):107-15.

131. Oettinger MA, Schatz DG, Gorka C, Baltimore D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science. 1990;248(4962):1517-23.

132. Hogquist KA, Baldwin TA, Jameson SC. Central tolerance: learning self-control in the thymus. Nat Rev Immunol. 2005;5(10):772-82.

133. Rowland SL, Tuttle K, Torres RM, Pelanda R. Antigen and cytokine receptor signals guide the development of the naive mature B cell repertoire. Immunol Res. 2013;55(1-3):231-40.

134. Pelanda R, Torres RM. Receptor editing for better or for worse. Curr Opin Immunol. 2006;18(2):184-90.

135. Hogquist KA, Jameson SC, Heath WR, Howard JL, Bevan MJ, Carbone FR. T cell receptor antagonist peptides induce positive selection. Cell. 1994;76(1):17-27.

136. Zehn D, Lee SY, Bevan MJ. Complete but curtailed T cell response to very low affinity antigen. Nature. 2009;458(7235):211-4.

137. Fu G, Vallee S, Rybakin V, McGuire MV, Ampudia J, Brockmeyer C, et al. Themis controls thymocyte selection through regulation of T cell antigen receptor-mediated signaling. Nat Immunol. 2009;10(8):848-56.

138. Lesourne R, Uehara S, Lee J, Song KD, Li L, Pinkhasov J, et al. THEMIS, a new T cell specific protein important for late thymocyte development. Nat Immunol. 2009;10(8):840-7.

139. Johnson AL, Aravind L, Shulzhenko N, Morgun A, Choi SY, Crockford TL, et al. Themis is a member of a new metazoan gene family and is required for completion of thymocyte positive selection. Nat Immunol. 2009;10(8):831-9.

140. Wang D, Zheng M, Lei L, Ji J, Yao Y, Qiu Y, et al. Tespa1 is involved in late thymocyte development through the regulation of TCR-mediated signaling. Nat Immunol. 2012;13(6):560-8.

141. Li QJ, Chau J, Ebert PJ, Sylvester G, Min H, Liu G, et al. miR-181a is an intrinsic modulator of T cell sensitivity and selection. Cell. 2007;129(1):147-61.

142. Fisher MR, Rivera-Reyes A, Bloch NB, Schatz DG, Bassing CH. Immature Lymphocytes Inhibit Rag1 and Rag2 Transcription and V(D)J Recombination in Response to DNA Double-Strand Breaks. J Immunol. 2017;198(7):2943-56.

143. Vettermann C, Schlissel MS. Allelic exclusion of immunoglobulin genes: models and mechanisms. Immunol Rev. 2010;237(1):22-42.

144. Peterson DO, Kriz KG, Marich JE, Toohey MG. Sequence organization and molecular cloning of mouse mammary tumor virus DNA endogenous to C57BL/6 mice. J Virol. 1985;54(2):525-31.

145. Salinas J, Zerial M, Filipski J, Crepin M, Bernardi G. Nonrandom distribution of MMTV proviral sequences in the mouse genome. Nucleic Acids Res. 1987;15(7):3009-22.

146. Palacios EH, Weiss A. Distinct roles for Syk and ZAP-70 during early thymocyte development. J Exp Med. 2007;204(7):1703-15.

147. Latour S, Chow LM, Veillette A. Differential intrinsic enzymatic activity of Syk and Zap-70 protein-tyrosine kinases. J Biol Chem. 1996;271(37):22782-90.

148. Steinel NC, Brady BL, Carpenter AC, Yang-Iot KS, Bassing CH. Posttranscriptional silencing of VbetaDJbetaCbeta genes contributes to TCRbeta allelic exclusion in mammalian lymphocytes. J Immunol. 2010;185(2):1055-62.

149. Germain RN, Stefanova I. The dynamics of T cell receptor signaling: complex orchestration and the key roles of tempo and cooperation. Annu Rev Immunol. 1999;17:467-522.

150. Corse E, Gottschalk RA, Allison JP. Strength of TCR–Peptide/MHC Interactions and In Vivo T Cell Responses. J Immunol. 2011;186(9):5039-45.

151. Van-Kaer L, Ashton-Rickardt PG, Ploegh HL, Tonegawa S. TAP1 mutant mice are deficient in antigen presentation, surface class I molecules, and CD4-8+ T cells. Cell. 1992;71(7):1205-14.

152. Franco S, Gostissa M, Zha S, Lombard DB, Murphy MM, Zarrin AA, et al. H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. Mol Cell. 2006;21(2):201-14.

153. Karo J, Schatz D, Sun J. The RAG Recombinase Dictates Functional Heterogeneity and Cellular Fitness in Natural Killer Cells. Cell. 2014;159(1):94-107.

154. Ostanin DV, Bao J, Koboziev I, Gray L, Robinson-Jackson SA, Kosloski-Davidson M, et al. T cell transfer model of chronic colitis: concepts, considerations, and tricks of the trade. Am J Physiol Gastrointest Liver Physiol. 2009;296(2):G135-46.

155. Eri R, McGuckin MA, Wadley R. T cell transfer model of colitis: a great tool to assess the contribution of T cells in chronic intestinal inflammation. Methods Mol Biol. 2012;844:261-75.

156. Mottet C, Uhlig HH, Powrie F. Cutting edge: cure of colitis by CD4+CD25+ regulatory T cells. J Immunol. 2003;170(8):3939-43.

157. Zimmermann J, Durek P, Kühl AA, Schattenberg F, Maschmeyer P, Siracusa F, et al. The intestinal microbiota determines the colitis-inducing potential of T-bet-deficient Th cells in mice. Eur J Immunol. 2018;48(1):161-7.

158. Reinoso-Webb C, den Bakker H, Koboziev I, Jones-Hall Y, Rao Kottapalli K, Ostanin D, et al. Differential Susceptibility to T Cell-Induced Colitis in Mice: Role of the Intestinal Microbiota. Inflamm Bowel Dis. 2018;24(2):361-79.

159. Ahern PP, Schiering C, Buonocore S, McGeachy MJ, Cua DJ, Maloy KJ, et al. Interleukin-23 Drives Intestinal Inflammation through Direct Activity on T Cells. Immunity. 2010;33(2-2):279-88.

160. Zimmermann J, Kuhl AA, Weber M, Grun JR, Loffler J, Haftmann C, et al. T-bet expression by Th cells promotes type 1 inflammation but is dispensable for colitis. Mucosal Immunol. 2016;9(6):1487-99.

161. Pieper K, Grimbacher B, Eibel H. B-cell biology and development. J Allergy Clin Immunol. 2013;131(4):959-71.

162. van-Zelm MC, Szczepański T, van der Burg M, van Dongen JJ. Replication history of B lymphocytes reveals homeostatic proliferation and extensive antigen-induced B cell expansion. J Exp Med. 2007;204(3):645-55.

163. Gayle MD, Sonya LS, Bart TE, Angela KD, Stephen CJ, Kristin AH. Preselection Thymocytes Are More Sensitive to T Cell Receptor Stimulation Than Mature T Cells. The Journal of Experimental Medicine. 1998;188(10):1867-74.

164. Peterson DA, DiPaolo RJ, Kanagawa O, Unanue ER. Cutting edge: negative selection of immature thymocytes by a few peptide-MHC complexes: differential sensitivity of immature and mature T cells. J Immunol. 1999;162(6):3117-20.

165. Brandle D, Muller S, Muller C, Hengartner H, Pircher H. Regulation of RAG-1 and CD69 expression in the thymus during positive and negative selection. Eur J Immunol. 1994;24(1):145-51.

166. Kuo TC, Schlissel MS. Mechanisms controlling expression of the RAG locus during lymphocyte development. Curr Opin Immunol. 2009;21(2):173-8.

167. Mingueneau M, Kreslavsky T, Gray D, Heng T, Cruse R, Ericson J, et al. The transcriptional landscape of αβ T cell differentiation. Nat Immunol. 2013;14(6):619-32.

168. Berkley AM, Fink PJ. CD8+ Recent Thymic Emigrants Exhibit Increased Responses to Low-Affinity Ligands and Improved Access to Peripheral Sites of Inflammation. J Immunol. 2014;93(7):3262-6.

169. Yu W, Nagaoka H, Jankovic M, Misulovin Z, Suh H, Rolink A, et al. Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. Nature. 1999;400(6745):682-7.

170. Mallaun M, Zenke G, Palmer E. A discrete affinity-driven elevation of ZAP-70 kinase activity initiates negative selection. J Recept Signal Transduct Res. 2010;30(6):430-43.

171. Sandberg JK, Chambers BJ, Van Kaer L, Karre K, Ljunggren HG. TAP1-deficient mice select a CD8+ T cell repertoire that displays both diversity and peptide specificity. Eur J Immunol. 1996;26(2):288-93.

172. Stepanek O, Prabhakar AS, Osswald C, King CG, Bulek A, Naeher D, et al. Co-receptor scanning by the T-cell receptor provides a mechanism for T cell tolerance. Cell. 2014;159(2):333-45.

173. Tang Q, Adams JY, Penaranda C, Melli K, Piaggio E, Sgouroudis E, et al. Central role of defective interleukin-2 production in the triggering of islet autoimmune destruction. Immunity. 2008;28(5):687-97.

174. Anderson MS, Bluestone JA. The NOD mouse: a model of immune dysregulation. Annu Rev Immunol. 2005;23:447-85.

175. Delovitch TL, Singh B. The nonobese diabetic mouse as a model of autoimmune diabetes: immune dysregulation gets the NOD. Immunity. 1997;7(6):727-38.

176. Moon JJ, Chu HH, Pepper M, McSorley SJ, Jameson SC, Kedl RM, et al. Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. Immunity. 2007;27(2):203-13.

177. Zhu J, Yamane H, Paul WE. Differentiation of Effector CD4 T Cell Populations*. Annu Rev Immunol. 2010;28:445-89.

178. Ji Y, Resch W, Corbett E, Yamane A, Casellas R, Schatz DG. The In Vivo Pattern of Binding of RAG1 and RAG2 to Antigen Receptor Loci. Cell. 2010;141(3):419-31.

179. Li L, Salido E, Zhou Y, Bhattacharyya S, Yannone SM, Dunn E, et al. Targeted disruption of the Artemis murine counterpart results in SCID and defective V(D)J recombination that is partially corrected with bone marrow transplantation. J Immunol. 2005;174(4):2420-8.

180. Xiao Z, Dunn E, Singh K, Khan IS, Yannone SM, Cowan MJ. A Non Leaky Artemis-Deficient Mouse that Accurately Models the Human SCID Phenotype Including Resistance to Hematopoietic Stem Cell Transplantation. Biol Blood Marrow Transplant. 2009;15(1):1-11.

181. Quadros RM, Miura H, Harms DW, Akatsuka H, Sato T, al e. Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucl... - PubMed - NCBI. Genome Biol. 2017;18(1).

182. Miura H, Quadros RM, Gurumurthy CB, Ohtsuka M. Easi-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors. Nat Protoc. 2018;13(1):195-215.

183. Balomenos D, Balderas RS, Mulvany KP, Kaye J, Kono DH, Theofilopoulos AN. Incomplete T cell receptor V beta allelic exclusion and dual V beta-expressing cells. J Immunol. 1995;155(7):3308-12.

184. von-Boehmer H, Melchers F. Checkpoints in lymphocyte development and autoimmune disease. Nat Immunol. 2010;11(1):14-20.

185. Niederberger N, Holmberg K, Alam SM, Sakati W, Naramura M, Gu H, et al. Allelic exclusion of the TCR alpha-chain is an active process requiring TCR-mediated signaling and c-Cbl. J Immunol. 2003;170(9):4557-63.

186. Walsh D, Mathews MB, Mohr I. Tinkering with Translation: Protein Synthesis in Virus-Infected Cells. Cold Spring Harb Perspect Biol. 52013.

187. Komar AA, Hatzoglou M. Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states. Cell Cycle. 102011. p. 229-40.

188. Hershey JW, Sonenberg N, Mathews MB. Principles of Translational Control: An Overview. Cold Spring Harb Perspect Biol. 42012.

189. Spriggs KA, Stoneley M, Bushell M, Willis AE. Re-programming of translation following cell stress allows IRES-mediated translation to predominate. Biol Cell. 2008;100(1):27-38.

190. Jopling CL, Spriggs KA, Mitchell SA, Stoneley M, Willis AE. L-Myc protein synthesis is initiated by internal ribosome entry. Rna. 2004;10(2):287-98.

191. Bullock TN, Eisenlohr LC. Ribosomal scanning past the primary initiation codon as a mechanism for expression of CTL epitopes encoded in alternative reading frames. J Exp Med. 1996;184(4):1319-29.

192. Babendure JR, Babendure JL, Ding JH, Tsien RY. Control of mammalian translation by mRNA structure near caps. RNA. 2006;12(5):851-61.

193. Pichon X, Wilson LA, Stoneley M, Bastide A, King HA, Somers J, et al. RNA binding protein/RNA element interactions and the control of translation. Curr Protein Pept Sci. 2012;13(4):294-304.

194. Romanelli MG, Diani E, Lievens PMJ. New Insights into Functional Roles of the Polypyrimidine Tract-Binding Protein. Int J Mol Sci. 142013. p. 22906-32.

195. King HA, Cobbold LC, Pichon X, Poyry T, Wilson LA, Booden H, et al. Remodelling of a polypyrimidine tract-binding protein complex during apoptosis activates cellular IRESs. Cell Death Differ. 2014;21(1):161-71.

196. Ingolia N, Lareau L, Weissman J. Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. Cell. 2011;147(4):789-802.

197. MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, et al. A systematic survey of loss-of-function variants in human protein-coding genes. Science. 2012;335(6070):823-8.

198. Chinn IK, Chan AY, Chen K, Chou J, Dorsey MJ, Hajjar J, et al. Diagnostic interpretation of genetic studies in patients with primary immunodeficiency diseases: A working group report of the Primary Immunodeficiency Diseases Committee of the American Academy of Allergy, Asthma & Immunology. J Allergy Clin Immunol. 2019.

199. Chavarria-Smith J, Vance RE. The NLRP1 inflammasomes. Immunol Rev. 2015;265(1):22-34.

200. Zhong FL, Mamai O, Sborgi L, Boussofara L, Hopkins R, Robinson K, et al. Germline NLRP1 Mutations Cause Skin Inflammatory and Cancer Susceptibility Syndromes via Inflammasome Activation. Cell. 2016;167(1):187-202.e17.

201. Malik A, Kanneganti TD. Inflammasome activation and assembly at a glance. J Cell Sci. 2017;130(23):3955-63.

202. Mao L, Kitani A, Strober W, Fuss IJ. The Role of NLRP3 and IL-1β in the Pathogenesis of Inflammatory Bowel Disease. Front Immunol. 2018;9.

203. Monteleone G, Trapasso F, Parrello T, Biancone L, Stella A, Iuliano R, et al. Bioactive IL-18 expression is up-regulated in Crohn's disease. J Immunol. 1999;163(1):143-7.

204. Williams MA, O'Callaghan A, Corr SC. IL-33 and IL-18 in Inflammatory Bowel Disease Etiology and Microbial Interactions. Front Immunol. 2019;10.

205. Kim H, Kim M, Im SK, Fang S. Mouse Cre-LoxP system: general principles to determine tissue-specific roles of target genes. Lab Anim Res. 2018;34(4):147-59.

206. Arbuckle JL, Rahman NS, Zhao S, Rodgers W, Rodgers KK. Elucidating the domain architecture and functions of non-core RAG1: the capacity of a non-core zinc-binding domain to function in nuclear import and nucleic acid binding. BMC Biochem. 2011;12:23.

207. Spanopoulou E, Cortes P, Shih C, Huang CM, Silver DP, Svec P, et al. Localization, interaction, and RNA binding properties of the V(D)J recombination-activating proteins RAG1 and RAG2. Immunity. 1995;3(6):715-26.

208. Maman Y, Teng G, Seth R, Kleinstein SH, Schatz DG. RAG1 targeting in the genome is dominated by chromatin interactions mediated by the non-core regions of RAG1 and RAG2. Nucleic Acids Res. 2016;44(20):9624-37.

209. Kannan M, Lee SJ, Schwedhelm-Domeyer N, Stegmuller J. The E3 ligase Cdh1-anaphase promoting complex operates upstream of the E3 ligase Smurf1 in the control of axon growth. Development. 2012;139(19):3600-12.

210. Wagner SA, Beli P, Weinert BT, Nielsen ML, Cox J, Mann M, et al. A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. Mol Cell Proteomics. 2011;10(10):M111 013284.

211. Udeshi ND, Mertins P, Svinkina T, Carr SA. Large-scale identification of ubiquitination sites by mass spectrometry. Nat Protoc. 2013;8(10):1950-60.

212. Udeshi ND, Svinkina T, Mertins P, Kuhn E, Mani DR, Qiao JW, et al. Refined preparation and use of anti-diglycine remnant (K-epsilon-GG) antibody enables routine quantification of 10,000s of ubiquitination sites in single proteomics experiments. Mol Cell Proteomics. 2013;12(3):825-31.

213. Brisse E, Wouters C, H., Matthys P. Hemophagocytic lymphohistiocytosis (HLH): A heterogeneous spectrum of cytokine-driven immune disorders. Cytokine & Growth Factor Reviews. 2015;26(3):263280.

214. Jordan M, B., Hildeman D, Kappler J, Marrack P. An animal model of hemophagocytic lymphohistiocytosis (HLH): CD8+ T cells and interferon gamma are essential for the disorder. Blood. 2004;104(3):735-43.

215. Rood JE, Rao S, Paessler M, Kreiger PA, Chu N, Stelekati E, et al. ST2 contributes to T-cell hyperactivation and fatal hemophagocytic lymphohistiocytosis in mice. Blood. 2016;127(4):426-35.

216. Taylor MD, Burn TN, Wherry EJ, Behrens EM. CD8 T Cell Memory Increases Immunopathology in the Perforin-Deficient Model of Hemophagocytic Lymphohistiocytosis Secondary to TNF-α. 2018.

217. Rood JE, Burn TN, Neal V, Chu N, Behrens EM. Disruption of IL-33 Signaling Limits Early CD8+ T Cell Effector Function Leading to Exhaustion in Murine Hemophagocytic Lymphohistiocytosis. Front Immunol. 2018;9.

218. Tesi B, Sieni E, Neves C, Romano F, Cetica V, Cordeiro A, I., et al. Hemophagocytic lymphohistiocytosis in 2 patients with underlying IFN-γ receptor deficiency. The Journal of allergy and clinical immunology. 2015;135(6):1638-41.

219. Burn TN, Weaver L, Rood JE, Chu N, Bodansky A, Kreiger PA, et al. Genetic Deficiency of Interferon-gamma Reveals Interferon-gamma-Independent Manifestations of Murine Hemophagocytic Lymphohistiocytosis. Arthritis Rheumatol. 2019.

220. Gopalakrishnan S ea. Unifying model for molecular determinants of the preselection Vβ repertoire. PNAS. 2013;110(34):E3206-15.

221. Shih HY, Verma-Gaur J, Torkamani A, Feeney AJ, Galjart N, Krangel MS. Tcra gene recombination is supported by a Tcra enhancer- and CTCF-dependent chromatin hub. Proc Natl Acad Sci U S A. 2012;109(50):E3493-502.

222. Chen L, Carico Z, Shih HY, Krangel MS. A discrete chromatin loop in the mouse Tcra-Tcrd locus shapes the TCRdelta and TCRalpha repertoires. Nat Immunol. 2015;16(10):1085-93.

223. Niiya M ea. Correction of ADAMTS13 deficiency by in utero gene transfer of lentiviral vector encoding ADAMTS13 genes. Mol Ther. 2009;17(1):34-41.

224. Zhang YH, Shetty K, Surleac MD, Petrescu AJ, Schatz DG. Mapping and Quantitation of the Interaction between the Recombination Activating Gene Proteins RAG1 and RAG2. J Biol Chem. 2015;290(19):11802-17.

225. Otsu M, Steinberg M, Ferrand C, Merida P, Rebouissou C, Tiberghien P, et al. Reconstitution of lymphoid development and function in ZAP-70-deficient mice following gene transfer into bone marrow cells. Blood. 2002;100(4):1248-56.