Identification of an inhibitory domain in RAG-2 that gates access of the V(D)J recombinase to chromatin

by
Alyssa Maria Ward

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
March, 2017

Abstract

V(D)J recombination is initiated by a specialized transposase consisting of the subunits RAG-1 and RAG-2. The susceptibility of gene segments to DNA cleavage by the V(D)J recombinase is correlated with epigenetic modifications characteristic of active chromatin, including trimethylation of histone H3 on lysine 4 (H3K4me3). Engagement of H3K4me3 by a plant homeodomain (PHD) in RAG-2 promotes recombination *in vivo* and stimulates DNA cleavage by RAG *in vitro*. We characterized features of this PHD-mediated inhibitory domain and identified a second, independent, inhibitory domain. Mutation of this second inhibitory domain allows bypass of the requirement for engagement of H3K4me3 by the RAG-2 PHD. Disruption of this inhibitory domain was associated with constitutive increases in recombination frequency, DNA cleavage activity, substrate binding affinity, and catalytic rate. We further characterized this domain genetically and determined that the inhibitory function is imposed by acidic residues between residues 352 and 405. Further, we were able to demonstrate that the inhibitory domain mutation is refractory to removal of the entire noncore portion of RAG-2. Therefore, this inhibitory domain acts independently of the PHD. Inactivation of the inhibitory domain confers a gain-of-function recombination phenotype and permits rearrangement at endogenous IgH and Igκ loci in the absence of H3K4me3 binding. In B progenitor cells, localization of wild-type RAG-2 to the IgH locus and actively transcribed loci is abolished by the W453A mutation, indicating that this pattern of chromatin localization is dependent on recognition of H3K4me3. This same mutation also abolishes association of RAG-1 with the IgH locus. Strikingly, disruption of the

inhibitory domain permits association of RAG-2 and RAG-1 with the IgH locus even in the absence of H3K4me3 engagement by RAG-2. Thus, the RAG-2 inhibitory domain serves as a binary gate that permits the association of RAG-1 and RAG-2 with chromatin only if H3K4me3 is engaged by the RAG-2 PHD finger.

Ph.D. Dissertation referees for Alyssa Maria Ward

Stephen Desiderio, M.D., Ph.D.

Director, Institute for Basic Biomedical Sciences

Professor, Molecular Biology and Genetics

Johns Hopkins University School of Medicine

(Thesis advisor)

Ranjan Sen, Ph.D.

Chief, Laboratory of Cellular and Molecular Biology

National Institute on Aging

(Reader)

Acknowledgments

I want to express my sincere gratitude for my advisor, Dr. Stephen Desiderio. He is not only a brilliant scientist, but one who genuinely cares about both the science and the students in the lab. Over the years, I've had triumphs and problems both within and outside of the lab. In every situation, I've known that Dr. Desiderio would be there with sympathy, support, or solutions. He listens to my ideas and always gives me the opportunity to be wrong, which has fostered my critical thinking and analysis skills. I have so much respect for Dr. Desiderio as a scientist, mentor, and as a person. I am thankful that I ended up in his lab, as I doubt I would have grown so much if I were in a different environment.

I am grateful for my colleagues in the Desiderio lab. Nick Dordai is more than a lab manager, he is a world-class problem solver who has helped me with more malfunctions than any one person should have. John Bettridge taught me how to perform biochemistry, but more importantly, our shared interests outside of science have helped to make science much less arduous. Similarly, Meiling May has been so much fun to have in lab and I am grateful for her help. Wenyan Lu is a supportive and helpful colleague. While not a part of the Desiderio lab, Shasha Wang has always been there for me when things got rocky. I'd also like to thank all the past members of the Desiderio lab. Chao Lu was an accomplished biochemist and a great person to be around. Christopher Cooper and Jihyun Lee both helped me immensely at the beginning of my studies and Taylor Reynolds was also there for me.

I would also like to thank my collaborators, Dr. Ranjan Sen and his post-doctoral fellow, Dr. Gita Kumari for their help in this work. Dr. Sen was also a member of my thesis committee, providing guidance and support throughout this project. I would also like to acknowledge Dr. Joel Pomerantz for his help serving on my thesis committee, but also always providing advice and useful feedback in passing. I would also like to thank Dr. Kathy Burns for serving on my thesis committee. I am grateful to have been a part of the Molecular Biology and Genetics and Immunology departments, where the faculty and students have provided helpful feedback for my project and my individual development.

Finally, I would like to thank my friends and family for their support throughout this endeavor. I did not know what to expect when I moved here, but I didn't expect to find anything as amazing as the supportive and fun people I've found here.

Table of Contents

List of Figures and Tables

Introduction

Adaptive immunity and receptor rearrangement

The adaptive immune system is characterized by its ability to specifically recognize a diverse spectrum of epitopes. Separately encoding this diversity would require a trillion base pairs, three times the number in the entire haploid mouse genome. The adaptive immune system solves this coding paradox by assembling antigen receptors from discrete DNA segments. For example, at the immunoglobulin heavy chain (IgH) locus in mice, there are approximately 110 functional variable (V) segments (Johnston et al. 2006), 10-13 diversity (D) segments (Ye 2004), and 4 joining (J) segments. These gene segments are each flanked by at least one recombination signal sequence (RSS) that is recognized by a complex of recombination activating genes 1 and 2 (RAG-1 and RAG-2). RSSs are composed of a conserved nonamer and heptamer separated by a spacer of 12 bp or 23 bp, termed 12-RSS and 23-RSS respectively. The RAG recombinase participates in all stages of V(D)J recombination, from RSS binding and synapsis of the gene segments, to cleavage and repair of the resulting double-strand breaks (DSBs). Expression of RAG-1 and RAG-2 is restricted to developing lymphocytes, however regulation of the antigen receptor loci prevents inappropriate recombination, even in the presence of RAG (Stanhope-Baker et al. 1996).

Antibodies, generated by B-cells, are composed of four polypeptides: two identical heavy chains and two identical light chains encoded by either the immunoglobulin κ (Igκ) or immunoglobulin λ (Igλ) loci. Disulfide bonds anchor the light chains to the heavy chains and connect the heavy chains to each other to form the characteristic "Y" structure of antibodies. T-cell receptors (TCRs) are heterodimers formed by recombination of the TCR α and β or ɣ and δ loci. B- and T-cells are clonal in

that each cell expresses only one of the billions of possible receptors. The importance of the adaptive immune system is illustrated in the disease phenotypes associated with a loss of V(D)J recombination. There is no redundancy in RAG function, therefore inactivating mutations result in a severe combined immunodeficiency (SCID) phenotype, marked by a complete lack of mature lymphocytes. Hypomorphic RAG alleles that retain some recombination activity result in Omenn syndrome, which is characterized by the absence of B-cells and oligoclonal, autoreactive T-cells (Villa et al. 2001).

Chromosomes are extended DNA molecules that must be accommodated in the relatively small nucleus of the cell in an ordered and dynamic structure that allows access to necessary information. This is accomplished by compacting DNA into nucleosomes. Each nucleosome consists of 146 bp of DNA wrapped around the histone core, composed of two molecules each of H2A, H2B, H3, and H4. Between these core nucleosome assemblies is linker DNA that can be further compacted by linker histones including H1. Nucleosomes and higher order structures can compact DNA more than 1,000-fold in dividing cells. Histones can be extensively modified and these modifications can be specifically recognized by a variety of protein domains.

The RAG recombinase

RAG-1 and RAG-2 form a "Y" shaped heterotetramer with two intertwined RAG-1 monomers forming the base and a RAG-2 monomor upon each (Kim et al. 2015; Grundy et al. 2009). RAG-1 is 1040 amino acids long, however only the region from 384 to 1008 is required for catalysis (Silver et al. 1993; Kirch et al. 1996; Sadofsky et al. 1993). The nonamer binding domain (NBD), from residue 391 to 459, binds the

conserved nonamer of the RSS. This domain adopts a three-helix structure that forms extensive hydrophobic contacts with another RAG-1 NBD to form a homodimer. The homodimer engages two RSS nonamers through extensive minor groove contacts, marginal major groove contacts, and interactions with the backbone of the spacer (Fang Yin et al. 2009; Kim et al. 2015). The two RAG-1 monomers are intertwined such that the NBD of one monomer interacts with the RSS that will be cleaved in the active site of the other monomer (Swanson 2001). A flexible linker connects the NBD to the dimerization and DNA binding domain (DDBD) that is highly positively charged and may, in collaboration with high mobility group box 1 (HMGB1) stabilize the kink in the RSS that facilitates cleavage (Kim et al. 2015). HMGB1 is a sequence non-specific DNA binding and bending protein that stimulates RAG binding to the 23-RSS and stabilizes DNA bending at the 12-RSS heptamer (Zagelbaum et al. 2016). RAG-1 adopts an RNase H fold that forms the catalytic site with D600, D708, and E962 (DDE motif). Mutation of the DDE motif abolishes catalytic activity while sparing RSS binding (Kim et al. 1999). Zinc is coordinated by two histidines and two cysteines. While core RAG-1 contains the portion of the protein minimally required for recombination, there is evidence that the noncore regions of RAG-1 contribute to recombination *in vivo* (Liang et al. 2002).

Concurrent investigations of the noncore regions of RAG-1 identified putative E3 ligase activity (Yurchenko et al. 2003; Jones & Gellert 2003) within the Really Interesting New Gene (RING) finger domain N-terminal of the RAG-1 core. In short, ubiquitination begins with a general ubiquitin-activating enzyme (E1) after which the activated ubiquitin is transferred to a ubiquitin-conjugating enzyme (E2) which works with the specific ubiquitin-protein ligase (E3) to attach ubiquitin to a lysine residue in the

target substrate (Laney & Hochstrasser 1999). RAG-dependent ubiquitin uptake has been observed in the presence of various E2 ligases (Yurchenko et al. 2003; Jones & Gellert 2003). The RAG-1 RING domain was shown to bind to (Cortes et al. 1994) and ubiquitinate (Simkus, Makiya, et al. 2009) karyopherin alpha 1 (KPNA1), an importin subunit. There is additional evidence of autoubiquitination (Jones & Gellert 2003; Singh & Gellert 2015) and RAG-dependent histone H3 ubiquitination (Grazini et al. 2010; Jones et al. 2011; Deng et al. 2015). Interestingly, a complex of proteins that co-purifies with full length RAG-1, termed the VDCR complex (VprBP, DDB1, Cul4A, and Roc1) exhibited ubiquitination activity that was stimulated by RAG-1 and conditional VprBP deletion resulted in recombination defects (Kassmeier et al. 2012). Purified, ubiquitinated RAG-1 exhibits increased basal activity (Singh & Gellert 2015) and mutations that impaired E3 ligase activity were associated with decreased recombination frequency (Simkus, Bhattacharyya, et al. 2009). There are also cases of Omenn syndrome resulting from mutations in the RING domain (Villa et al. 2001; Simkus et al. 2007; Deng et al. 2015). Taken together, these data support a role for noncore RAG-1 in recombination, however the nature of this role is still unclear.

RAG-2

RAG-2 is 527 amino acid residues long. Only the N-terminal 387 amino acids, termed the core region, are required for RSS cleavage *in vitro* (Sadofsky et al. 1994; Cuomo & Oettinger 1994), but removal of the noncore region, comprising residues 388 through 527, is associated with decreased recombination frequency (Steen et al. 1999) and increased aberrant recombination *in vivo* (Sekiguchi et al. 2001; Akamatsu et al.

2003). Core RAG-2 consists of six kelch-like repeats that fold into a β-propeller structure with extended loops that interface with the zinc coordinating and RNase H domains of RAG-1 (Kim et al. 2015; Callebaut & Mornon 1998). These extensive contacts may explain the requirement of RAG-2 for efficient recombination despite RAG-1 containing the important catalytic and DNA binding residues. The noncore region of RAG-2 serves several regulatory functions, including cell cycle-dependent degradation (Jiang et al. 2005; Lee & Desiderio 1999; Li et al. 1996; Zhang et al. 2011; Lin & Desiderio 1994), nuclear import (Ross et al. 2003), recognition of H3K4me3 (Liu et al. 2007; Matthews et al. 2007; Ramón-Maiques et al. 2007), and autoinhibition (Lu et al. 2015).

While RAG-1 protein levels are steady throughout the cell cycle, RAG-2 expression is regulated dynamically. Threonine 490 (T490) is phosphorylated by cyclin A/CDK2 (Lin & Desiderio 1993), and this phosphorylation leads to ubiquitination by the Skp2-SCF E3 ligase. Polyubiquitination of RAG-2 leads to proteasomal degradation (Jiang et al. 2005) before cells enter S phase, which restricts recombination to G_0/G_1 (Lin & Desiderio 1994). RAG-2 bearing the point mutation, T490A, persists throughout the cell cycle (Li et al. 1996), and is associated with genomic instability and, in p53-deficient mice, lymphomagenesis (Zhang et al. 2011). This was recapitulated with core RAG-2 mice on a p53-null background (Chaumeil et al. 2013), however the noncore region serves multiple functions that confound analyses of core RAG-2 mice. The transcription factor, p53, is phosphorylated by proteins involved in repair of RAG-mediated breaks and its activation leads to cell cycle arrest, preventing cell cycle progression until RAG-mediated breaks have been repaired (Helmink & Sleckman 2012). Multiple mechanisms

ensure that RAG-mediated breaks are only generated in G_0/G_1 and resolved prior to initiation of S-phase.

RAG-2 contains a plant homeodomain (PHD), spanning residues 415 through 487, that binds histone H3K4me3 (Liu et al. 2007; Matthews et al. 2007; Ramón-Maiques et al. 2007). The RAG-2 PHD has a hydrophobic channel that accommodates the H3K4me3 side chain, coordinating the trimethyl ammonium with the conserved residue tryptophan 453 (W453) (Ramón-Maiques et al. 2007). A point mutation of W453 ablates the ability to bind to H3K4me3 (Ramón-Maiques et al. 2007; Matthews et al. 2007; Liu et al. 2007), impairing recombination on chromatinized substrates and at endogenous loci while sparing *in vitro* activity (Matthews et al. 2007; Liu et al. 2007). RAG-2 localization appears to be mediated largely or wholly by the PHD, as RAG-2 binds broadly across the genome in a pattern that is correlated with sites of H3K4me3 deposition. Further, this localization is independent of RAG-1 coexpression. RAG-1, by contrast, binds specifically to antigen receptor loci and can bind to all but the IgH locus in the absence of RAG-2 (Ji et al. 2010). A patient with Omenn syndrome was identified with a W453R mutation, underscoring the importance of the PHD function to recombination (Gomez et al. 2000).

Mechanism of V(D)J recombination

There are two classes of RSS, termed 12-RSS and 23-RSS, composed of conserved nonamer (5'-GGTTTTTGT) and heptamer (5'-CACAGTG) elements (Sakano et al. 1979) separated by spacers of 12 bp or 23 bp, respectively (Akira et al. 1987). Rearrangement proceeds in the following sequence: (1) capture by RAG of a 12-RSS and

a 23-RSS, resulting in synapsis of participating gene segments; (2) nicking by RAG at the junction between each gene segment and its flanking RSS; (3) transesterification to produce double-strand breaks; and (4) joining of the participating gene segments by classical non-homologous end joining (cNHEJ) (Gellert 2002; Deriano & Roth 2013). It is still unclear whether RAG-1 and RAG-2 form a complex that scans for appropriate substrates or if the RAG complex is assembled at the target RSS (Askary et al. 2014). RAG-1 does bind to most antigen receptor loci in the absence of RAG-2 and RAG-2 binds promiscuously to genomic H3K4me3 in a RAG-1-independent manner (Ji et al. 2010). Therefore, a parallel binding model must also account for the individual binding activities of the two RAG subunits. While RAG is capable of nicking at a single RSS *in vitro* (Yu & Lieber 2000), modeling suggests that synapsis precedes nicking and that synaptic complex formation is essentially irreversible (Askary et al. 2014). The cleavage mechanism is a two-step process whereby RAG first induces a single strand nick at the 5' end of the RSS, between the heptamer and coding sequence. The free hydroxyl attacks the opposite strand to create a hairpin coding end and a blunt 5' phosphorylated signal end (McBlane et al. 1995). Breaks are then resolved in collaboration with cNHEJ machinery.

Heptamer and nonamer sequences of the RSSs are well conserved and some mutations are prohibitive to hairpin formation. This can be ameliorated by base-unpairing, which suggests that the ability of RAG to unpair the sequence, rather than the identity of the sequence, determines recombination competency. There are distinct requirements for the different steps of catalysis. RAG binding is largely guided by the nonamer sequence, however nicking and hairpin formation rely primarily on the

heptamer. Nick generation is not dependent on the first two nucleotides of the heptamer, however hairpin formation is. Modeling shows that RAG can discriminate between appropriate recombination targets, those with an intact RSS and permissive euchromatin, and inappropriate targets lacking either of these features (Askary et al. 2014). This fidelity of substrate recognition is vital to maintaining genomic integrity.

The 12/23 rule

RAG exhibits a 50-fold preference for a synaptic complex containing one 12-RSS and one 23-RSS, a generalization termed the 12/23 rule. Enforcement of the 12/23 rule helps to prevent homotypic recombination amongst two of the same gene segment. The difference between a 12- and 23-RSS is approximately one helical turn of B-form DNA. This opens the possibility that the 12/23 rule is enforced, in part, by a helical phase requirement for RAG activity. This is supported by the finding that increasing the spacer length by one half-turn is more detrimental than increasing the spacer length by a whole helical turn (Ramsden et al. 1996). *In vitro* 12/23 specificity is improved by addition of HMGB1, which also stimulates recombination (van Gent et al. 1997). Free RAG heterotetramers are symmetric (Kim et al. 2015), however binding a 12- and 23-RSS will inherently induce asymmetry because the dimerized NBD domains both tilt towards the shorter 12-RSS. Cryo-electron microscopy experiments suggest that this asymmetry in RSS binding enforces the 12/23 rule, because a shorter 12-RSS means that a longer 23-RSS is required to make up for the tilt (Ru et al. 2015). The role of HMGB1 in creation and maintenance of these DNA bends, provides a molecular hypothesis explaining how HMGB1 aids in enforcement of the 12/23 rule. Further, stabilization of the bent DNA

conformation can mechanistically explain the effect of HMGB1 on the catalytic rate of the V(D)J recombinase.

While permitted by the 12/23 rule, V segments at the IgH and TCRβ loci do not typically recombine to J segments. This restriction, termed beyond 12/23, likely ensures that D segments are used. At the TCRβ locus, it was found that the Dβ1 12-RSS is inherently predisposed to recombine with the Vβ 23-RSS. Similarly, the Vβ 23-RSS would not recombine with Jβ 12-RSS. Beyond 12/23 regulation appears to be a sequence-dependent, position-independent phenomena (Bassing et al. 2000) that ensures each chain includes the appropriate gene segments.

Classical repair of RAG-mediated breaks

The hairpin coding ends and blunt signal ends formed by RAG-mediated cleavage remain stably associated with RAG in a post-cleavage complex. This is hypothesized to ensure that RAG-mediated breaks are funneled into the cNHEJ repair pathway. The Ku heterodimer, composed of Ku70 and Ku80 subunits, is highly abundant in cells and exhibits a high affinity for DNA breaks, including blunt ends, overhangs, and covalently sealed hairpins. Evidence suggests that Ku associates with noncore RAG-1, which may be one of the links between RAG cleavage and repair (Raval et al. 2008). Ku forms a ring that encircles DNA and translocates on the DNA strand to allow more Ku dimers to bind and coat DNA ends. Ku recruits DNA-PK_{CS} and further translocates to open up DNA ends to DNA-PK_{CS}. DNA-PK_{CS} bound to Ku and DNA ends is considered a holoenzyme that can then phosphorylate itself and many targets involved in DSB repair (Deriano & Roth 2013). While many substrates have been identified, including Artemis, Ku, the

histone variant H2AX, XRCC4, Ligase IV, and XLF, individual mutational analyses reveal redundancy. The Ku complex is required for efficient resolution of both coding and signal ends. While DNA-PK_{CS} is essential for coding joint formation, signal joint formation is only mildly impaired by its absence. This is likely due to functional redundancy with ataxia telangiectasia mutated (ATM), as double mutants have severe defects in signal joint formation (Gapud & Sleckman 2011). Artemis has basal 5' to 3' exonuclease activity, however DNA-PK_{CS} activates endonucleolytic activity that enables Artemis to hydrolyze the phosphodiester bonds covalently sealing the coding hairpins. Opening at the apex could lead to a flush coding joint with no gain or loss of nucleotides. Conversely, an asymmetric nick results in palindromic (P) nucleotides. The RAG complex can nick hairpin coding joints via sequence non-specific endonuclease activity that uses the same active site as recombination. This activity is associated with formation of coding joints with long P insertions (Besmer et al. 1998; Shockett & Schatz 1999).

Efficient repair of both signal and coding ends requires XRCC4 and DNA Ligase IV. DNA Ligase IV, in complex with XRCC4 and XLF/Cernunnos, ligates RAG-mediated breaks. XRCC4 stabilizes DNA Ligase IV, stimulates the first step in ligation, and may align DNA ends (Helmink & Sleckman 2012). Further, XRCC4 interacts with polynucleotide kinase/phosphatase which promotes DNA ligation by phosphorylating 5'-hydroxyl groups and dephosphorylating 3' phosphates to generate compatible ends. XLF and XRCC4 form filamentous structures that could function to bridge DSB gaps (Deriano & Roth 2013). MRE11, Rad50 and Nbs1 form the MRN complex, which activates ATM, initiates end resection, and tethers broken DNA ends. Once ATM is activated, it phosphorylates factors involved in DSB repair, including H2AX. ATM also

phosphorylates p53, which ties cell cycle progression to resolution of RAG-mediated DSBs. The phosphorylated form of histone variant H2AX, termed ɣ-H2AX, protects RAG induced DSBs from extensive end resection. Loss of ɣ-H2AX does not impair recombination, likely due to redundancy of some functions shared with XLF.

Terminal deoxynucleotidyltransferase (TdT) adds (N) nucleotides to 3' DNA ends in a template-independent manner. In V(D)J recombination, these N additions on coding ends increase junctional diversity (Helmink & Sleckman 2012). N and P nucleotides increase the variability of antigen receptors, even when the same gene segments are used. Signal ends are generally ligated without the additional processing steps, creating a flush joint between the heptamers of the 12- and 23-RSSs.

Alternative repair mechanisms

While RAG-mediated breaks are generally repaired in collaboration with the cNHEJ machinery, some evidence suggests that impairment of cNHEJ funnels DSBs into a distinct pathway called alternative non-homologous end joining (altNHEJ). Junctions repaired by altNHEJ frequently exhibit large deletions, microhomology, and large insertions of unknown origin. While relatively little is known about altNHEJ, it is thought to begin with extensive end resection and end with ligation by DNA Ligase III. In altNHEJ, Poly(ADP-ribose) polymerase 1 (PARP1) binds DNA analogously to and in competition with Ku. MRE11 has been implicated in the extensive end resection, and CtIP can open coding end hairpins in the absence of Artemis. altNHEJ may function in cNHEJ-sufficient cells as well, operating at one tenth the frequency of cNHEJ. It is difficult to determine whether altNHEJ represents a distinct pathway or functional

redundancy in cNHEJ. The features of junctions formed by altNHEJ are similar even when the deficiencies are in different parts of cNHEJ, which provides some evidence that altNHEJ represents a distinct pathway (Deriano & Roth 2013). However, the features that characterize altNHEJ are not unique, as cNHEJ frequently uses 1-2 bp of microhomology (Pannunzio et al. 2014). Further work will be required to determine whether cNHEJ adapts to deficiencies or if cNHEJ deficiencies funnel DSBs to a distinct, more error prone, pathway.

It is currently unclear what role the RAG complex plays in repair during V(D)J recombination. XLF is not strictly required for V(D)J recombination, however its deletion impairs cNHEJ in other contexts. This suggests that the function of XLF is partially redundant with a factor specific to V(D)J recombination. Both XLF null and core RAG-2 mice exhibit only mild reduction in T lymphopoesis, however crossing the two lines exacerbates this phenotype. The double mutant mice have impaired V(D)J recombination and resultant lymphopenia. This implicates the noncore region of RAG-2 in the ability to bypass XLF-deficiency, suggesting functional redundancy (Lescale et al. 2016). A RAG-2 frameshift mutation, RAG-2(FS361), results in 28 novel amino acids before a stop codon and elimination of the noncore regions of RAG-2. Mice expressing this variant are predisposed to lymphoma on a p53-null background, consistent with the evidence that cell-cycle regulation in noncore RAG-2 prevents genomic instability. However this mutation also allowed bypass of DNA-PK_{CS} and XRCC4 deficiencies, with junctions that bear hallmarks of altNHEJ (Corneo et al. 2007). Both this frame shift mutation and a truncation of RAG at 352 increased aberrant recombination and allowed bypass of Ku-deficiency (Gigi et al. 2014). Further, mutation or deletion of the acidic

hinge region of RAG-2 may predispose cells to use of altNHEJ, allowing bypass of deficiencies in classical NHEJ. Neutralization or deletion of the acidic hinge increased the appearance of aberrant metaphases, suggesting that the region is important to proper repair of RAG mediated lesions (Coussens et al. 2013). This acidic hinge region is similar to the inhibitory domain identified in studies below that can bypass the detrimental W453A mutation (Lu et al. 2015), which opens the possibility that the region could be important to both cleavage and repair. Taken together, these data implicate the C-terminus of RAG-2 in repair of RAG-mediated breaks through a not yet understood mechanism.

Cell-cycle dependent regulation of RAG-2 protein accumulation has implications on the mode of coding end repair. cNHEJ is active throughout the cell cycle, however homologous recombination (HR) predominates after S-phase, when a homologous chromosome is available for template-dependent repair. Mistimed repair of RAG-mediated breaks is associated with complex chromosomal translocations (Zhang et al. 2011) consistent with abortive iterations of break induced replication. Break induced replication in eukaryotic cells is not well studied, but involves strand invasion followed by establishment of an illicit replication fork (Malkova & Ira 2013). The result of this replication depends on the template used for repair: sister chromatids result in faithful repair; homologous chromosomes result in a loss of heterozygosity; and heterologous chromosomes result in non-reciprocal translocation. Tandem duplications arise in junctions of RAG-2(T490A) mice, which implies the use of a replicative repair mechanism. Thus, in addition to genetic interactions with repair factors, the noncore region of RAG-2 enforces timing of RAG activity that prevents involvement of HR.

Developmental regulation of V(D)J recombination

While a 12-RSS and 23-RSS are sufficient to support DNA cleavage by RAG *in vitro*, V(D)J recombination *in vivo* is subject to higher level constraints that restrict rearrangement to particular sets of gene segments in distinct lymphocyte lineages and developmental stages. Addition of RAG complexes to isolated nuclei from a given lymphoid cell type results in recombination of only the developmentally appropriate receptor (Stanhope-Baker et al. 1996). That is, the context of the antigen receptor loci determines which will recombine. In early B-cell development, the IgH and Igκ loci move from the nuclear periphery to a more central and accessible location (Kosak et al. 2002). Rearrangement is initiated first at the IgH locus. Within the IgH locus, D_H-J_H joining occurs on both alleles (Coleclough et al. 1981) and precedes recombination of a V_H segment to the D_H-J_H unit (Alt et al. 1984). The cell initially expresses this V(D)J product with a pair of surrogate light chains. Approximately two-thirds of the rearrangements are unproductive (Coleclough et al. 1981), in which case the second allele is rearranged. Upon productive rearrangement and expression of immunoglobulin μ chain, further recombination at the IgH locus is suppressed and rearrangement of light chain loci is initiated (Jung et al. 2006). This transition is associated with epigenetic changes at the Igκ locus and delayed changes at the Igλ locus, providing a plausible explanation for the observation that Igκ generally recombines prior to Igλ (Goldmit et al. 2005; Xu & Feeney 2009). Recombination is an ordered process that is highly regulated during lymphocyte development.

V(D)J recombination is invariably preceded by sterile germline transcription from promoters whose activity is positively correlated with rearrangement (Van Ness et al.

1981; Yancopoulos & Alt 1985). At the IgH locus, for example, μ0 transcripts initiate at the DQ52 promoter (Alessandrini & Desiderio 1991), and Iμ transcripts originate within the Eμ enhancer (Lennon & Perry 1985; Su & Kadesch 1990). Activation of germline transcription at antigen receptor loci is accompanied by alterations in chromatin organization and the establishment of chemical modifications characteristic of active chromatin, including H3K4me3 (Chakraborty et al. 2007; Goldmit et al. 2005; Liu et al. 2007; Matthews et al. 2007; Morshead et al. 2003; Subrahmanyam et al. 2012). It is not possible to uncouple epigenetic changes from germline transcription, which leaves at least three basic hypotheses about the relationship between transcription and epigenetic changes: (1) Epigenetic changes at active loci induce transcription; (2) Transcription at active loci induces epigenetic changes; or (3) Both epigenetic changes and transcription are secondary to another process. Emerging evidence shows that some epigenetic changes are important beyond any potential role in facilitating transcription. For example, DNA demethylation was observed at junctions at the IgH locus (Selimyan et al. 2013). The specificity for junctions may indicate that DNA demethylation marks segments for recombination. Similarly, monoallelic Igκ demethylation was observed and proposed to be a mechanism of maintaining allelic exclusion (Mostoslavsky et al. 1998). Allelic exclusion, or expression of only a single receptor, is critical to maintaining the monospecificity of lymphoid cells. There are likely multiple mechanisms that work in concert to enforce allelic exclusion. One proposed mechanism is a stochastic mechanism based on the observation that most rearrangements are unproductive, therefore simultaneous recombination would be unlikely to result in two productive rearrangements (Coleclough et al. 1981). After ATM recognizes DSBs in immature lymphocytes, a

signaling pathway is initiated to decrease RAG expression (Fisher et al. 2017), decreasing the probability of further recombination. Productive rearrangement also signals RAG down-regulation. Many mechanisms work in unison to ensure recombination occurs only at appropriate loci and that recombination ceases after a single productive rearrangement.

Epigenetic regulation of RAG activity

The recognition of H3K4me3 by the RAG-2 PHD is not merely a passive localization signal. Rather, engagement of H3K4me3 by the PHD finger stimulates recombination *in vivo* (Liu et al. 2007; Matthews et al. 2007) and cleavage of RSS substrates *in vitro* (Shimazaki et al. 2009; Grundy et al. 2010; Lu et al. 2015), conferring increases in substrate affinity and catalytic rate (Shimazaki et al. 2009; Lu et al. 2015). Affinity for substrate is also increased in the presence of H3K4me3 (Grundy et al. 2010; Lu et al. 2015). Evidence suggests that the RAG-2 PHD binds more strongly to H3K4me3 peptide with an additional symmetrical dimethylation at arginine 2 (H3R2me2s) (Ramón-Maiques et al. 2007). Additionally, H3R2me2s can be found *in vivo* and correlates with euchromatin. In pro-B cells, H3R2me2s is positively correlated with H3K4me3 across the genome, including the IgH locus (Yuan et al. 2012). Some evidence implicates the acidic region between the RAG-2 core and PHD in binding to histones. Further, mutations that decreased histone binding ability were associated with decreased recombination at endogenous loci (West et al. 2005). Additionally, the C-terminus of RAG-2 has also been proposed to bind to phosphoinositides (PI). Mutations that decreased PI binding also impaired recombination activity. However the W453R mutation of the PHD reduced PI binding despite the fact that PI binding was localized

outside of the PHD (Elkin et al. 2005). Therefore, H3K4me3 may not be the only chromatin feature recognized by the noncore RAG-2.

The role of histones in V(D)J recombination does not appear to be limited to interactions with histone tails. Nucleosome structures inhibit RAG activity *in vitro* (Du et al. 2008; Golding et al. 1999; Kwon et al. 2000; McBlane & Boyes 2000), ostensibly through denying access to the RSS. Nucleosome remodeling with the SWI/SNF ATP-dependent chromatin remodeling complex alleviated this repression (Du et al. 2008; Kwon et al. 2000). Hyperacetylation may bypass this inhibition in some contexts, however the data are equivocal. *In vitro*, hyperacetylation synergized with SWI/SNF activity to increase recombination (Kwon et al. 2000) and stimulated recombination activity on its own (McBlane & Boyes 2000). However, a report using similar methods provided evidence that neither HMGB1 nor hyperacetylation could bypass inhibition by nucleosomes (Golding et al. 1999). It is possible that variable results are due to differences in experimental conditions, so further work is necessary to elucidate the interaction between nucleosomes and acetylation *in vitro*. *In vivo*, histone acetylation is positively correlated with recombination (McMurry & Krangel 2000), however hyperacetylation is generally associated with active enhancers and transcription, both of which are intimately tied to V(D)J recombination. An analysis of RAG-1 binding in lymphoid cells determined that RAG-1 exhibits two binding modes, one that is explained by H3K4me3 and one that is driven by H3K27Ac, primarily in the absence of a functional RAG-2 PHD (Maman et al. 2016). The RING domain may mediate this interaction, as it has been proposed to bind to histone H3 (Deng et al. 2015; Grazini et al. 2010; Jones et al. 2011) and the H3K27Ac localization was dependent on noncore RAG-

1 (Maman et al. 2016). Additionally, polyubiquitination of RAG-1 has been suggested to increase recombination frequency. This mode of regulation operates independent of H3K4me3 stimulation, as recombination was higher with both H3K4me3 and ubiquitination than with either alone (Singh & Gellert 2015). However, RAG-1 isolated from human cells is not polyubiquitinated, so further work is necessary to determine whether stimulation of RAG activity by autoubiquitination has physiological relevance. While this bears further study, these data suggest that there is a role for noncore RAG-1 in regulation of recombination *in vivo*.

These direct chromatin interactions represent only one type of epigenetic regulation of recombination. As V(D)J recombination requires synapsis of gene segments over one mega base pair away. CCCTC-binding factor (CTCF) is a zinc-finger protein that binds CTCF-binding elements (CBEs) and has been implicated in long range chromosome interactions. The IgH locus contains multiple CBEs, the deletion of which changes VH usage (Lin et al. 2015). The current model is that DNA looping proteins like CTCF function to increase diversity in V(D)J recombination by bringing distal gene segments proximal (Choi & Feeney 2014; Medvedovic et al. 2013). These long-range interactions are also important in the epigenetic regulation of recombination.

Given that the V(D)J recombinase DNA-binding domains and catalytic core are largely contained within RAG-1 (Kim et al. 2015), the interactions between RAG-2 and H3K4me3 are consistent with the interpretation that H3K4me3 is an allosteric activator of the V(D)J recombinase. This interpretation was reinforced by the finding that binding of H3K4me3 to the RAG-2 PHD finger induces conformational changes in RAG-1 within the DNA-binding domains and a domain that acts as a scaffold for the catalytic core

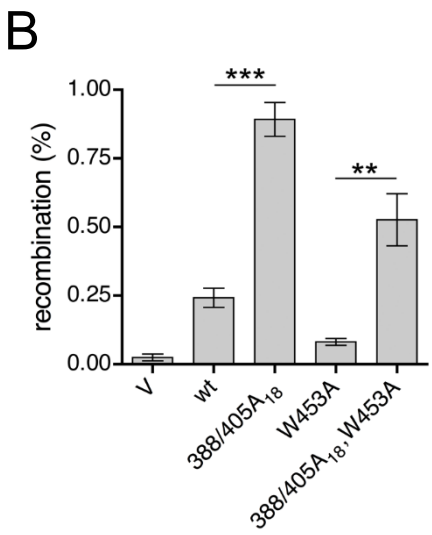
(Bettridge et al. 2017). The point mutation, W453A, impairs recombination more than removal of the noncore region. We therefore hypothesized that the PHD inhibits RAG activity and the binding of H3K4me3 relieves this inhibition. This model predicts the existence of a compensatory mutation in RAG that would render the complex impervious to inhibition by the PHD.

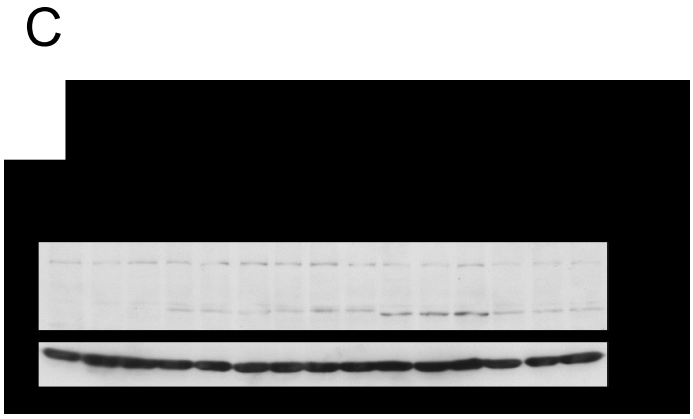
Results

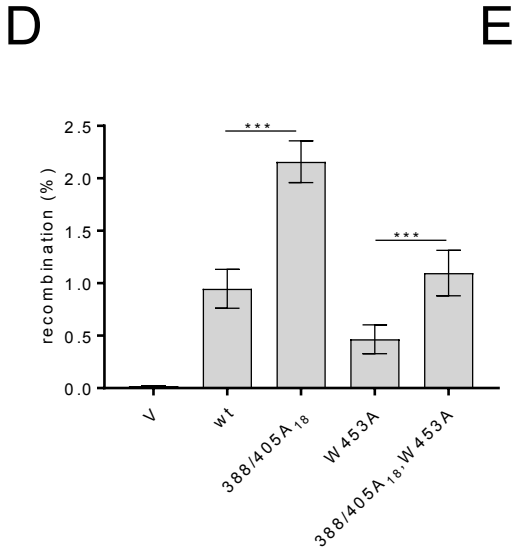
Identification of an inhibitory domain within the noncore region of RAG-2

That the deletion of the entire noncore region of RAG-2 would be less detrimental than a point mutation in the PHD suggested the presence of an inhibitory domain in RAG-2. We reasoned that this domain would be identifiable by second site mutations that rescue the activity of RAG-2(W453A). Therefore, we scanned the canonical noncore region of RAG-2 with clustered alanine substitutions of nine or ten amino acids each in the context of the loss of function RAG-2 W453A mutation. A preliminary screen identified two contiguous regions, 388-396 and 397-405, that rescued RAG-2(W453A) in an assay for V(D)J recombination (Liu 2009). Therefore, we constructed RAG-2 mutants bearing an alanine substitution spanning residues 388 through 405, RAG-2(388-405A_{18}), and the same clustered alanine mutation in the context of W453A, RAG-2(388-405A_{18}, W453A) (Figure 1A). Recombination activity was impaired in the RAG-2(W453A) mutant relative to wild type, as reported previously (Liu et al. 2007). In contrast, the double mutant, RAG-2(388-405A_{18}, W453A) was as active as wild-type. Moreover, the single mutant, RAG-2(388-405A_{18}) was significantly more active than wild-type (Figure 1B). Differences in protein accumulation did not account for the differences in recombination activity (Figure 1C). The gain of function mutation was consistently observed in distinct assays for signal joining (data not shown) and in an assay for coding joining (Figure 1D), as well as in combination with a mutation (T490A) that uncouples RAG-2 accumulation from the cell cycle (Figure 1E). Additionally, the 388-405A_{18} mutation exhibited no defect in precision of signal joining (Figure 2).


A
1
388
414
487
527
Core
L
PHD
D
334
456
DNKQAMSEAFYFYTLRCSEEDLSEDQKIVSNSQTSTEDPGDSTPFEDSEEFCFSAEATSFDGDDEFDTYNEDDEDDESVTGYWITCCPTCDVDINTWVPFYSTELNKPAMIYCSHGDGHWVHA Wildtype
RAG-2(W453A)
RAG-2(388/396A9,W453A)
RAG-2(397/405A9,W453A)
RAG-2(406/414A9,W453A)
RAG-2(388/405A18,W453A)
RAG-2(415/424A10)
RAG-2(406/414A9)
RAG-2(388/405A18)
RAG-2(370/387A18)
RAG-2(352/369A18)
RAG-2(334/351A18)


B
recombination (%)
1.00
0.75
0.50
0.25
0.00
***
**
V
wt
388/405A18
W453A
388/405A18,W453A


C


D
recombination (%)
2.5
2.0
1.5
1.0
0.5
0.0
***
***
V
wt
388/405A18
W453A
388/405A18,W453A

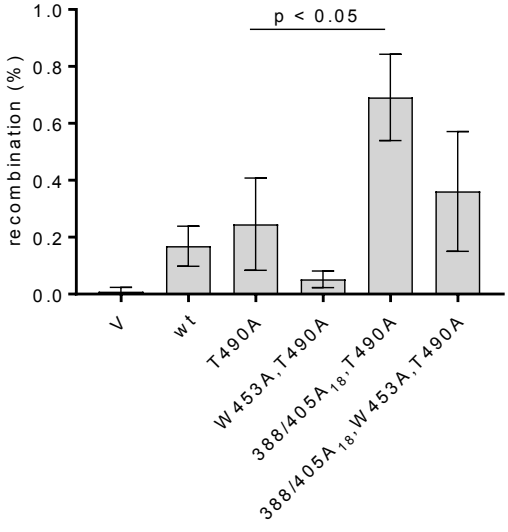

E
recombination (%)
1.0
0.8
0.6
0.4
0.2
0.0
p < 0.05
V
wt
T490A
W453A,T490A
388/405A18,T490A
388/405A18,W453A,T490A

Figure 1. The RAG-2 noncore region contains an inhibitory domain

(A) RAG-2 mutants. Above: diagram of RAG-2. Core refers to the canonical core region; linker (L), plant homeodomain (PHD), and signal for cell cycle-dependent degradation (D) are indicated. Initial and terminal amino acid residues as well as residues bounding the core, linker and PHD finger are numbered. Below: amino acid sequences (from residue 334 through 456) of wild-type RAG-2 and RAG-2 mutants. Red type indicates residues residing in the canonical core region. Hyphens indicate identity to wild-type.

(B) Rescue of an inactivating PHD finger mutation on signal joints. Recombination (%) is a calculation of the frequency of signal joining of the pJH200 signal joint recombination substrate (Hesse et al. 1987). Substrate, full-length RAG-1, and the indicated RAG-2 variant (vector (V), wild-type RAG-2 (wt), or indicated mutant) were transfected into NIH3T3 cells. The bars represent the mean ± SEM, n = 3 independent transfections, representing ≥ 500 ampicillin (A)-resistant colonies per RAG-2 variant and ≥ 200 A-resistant colonies for vector alone. Significant differences were determined by ANOVA (***$p < 0.001$; **$p < 0.01$).

(C) Immunodetection of RAG fusion proteins in recombination assays. NIH3T3 cells were co-transfected in triplicate with pJH200, RAG-1 and vector (pcDNA1), wild-type RAG-2, RAG-2(388/405A18), RAG-2(W453A) or RAG- 2(388/405A18, W453A). An aliquot of cells from each transfection was lysed and RAG proteins (upper panel) or actin (lower panel) were detected by immunoblotting with an anti-myc or anti-actin antibody, respectively. Positions of RAG-1 and RAG-2 fusion proteins are indicated.

(D) Rescue of an inactivating PHD finger mutation on coding joints. As in (B) performed with pJH290, a surrogate for coding joint formation.

(E) The RAG-2 388/405A18 mutation stimulates signal joining independent of cell cycle-dependent regulation of RAG activity. As in (B).

A

wt RAG-2

	12-RSS	23-RSS
4.1	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.2	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.3	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.4	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.5	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.6	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.7	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.8	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.9	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.10	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.11	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.12	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.13	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.14	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.15	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
4.16	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.1	----------	----------GTCTGGCTGTACNAAAACC
5.2	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.3	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.4	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.5	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.6	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.7	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.8	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.9	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.1	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.11	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.12	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.13	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.14	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.15	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
5.16	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.1	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.2	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.3	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.4	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.5	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.7	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.8	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.9	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.10	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.11	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.12	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.13	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.14	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.15	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
6.16	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC

B

RAG-2(388/405A$_{18}$)

	12-RSS	23-RSS
16.1	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.2	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.3	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.4	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.5	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.6	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.7	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.8	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.9	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.10	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.11	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.12	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.13	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.14	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.15	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
16.16	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.1	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.2	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.3	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.4	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.5	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.6	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.7	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.8	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.9	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.10	TGTTTTTGTTCCAGTCTGTAGCACTGTG C	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.11	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.12	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.13	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.14	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.15	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
17.16	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.1	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.2	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.3	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.4	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.5	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.6	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.7	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.8	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.9	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.10	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.11	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.12	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.13	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.14	TGTTTTTGTTCCAGTCTGTAGCACTGTG	-------GTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.15	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
18.16	TGTTTTTGTTCCAGTCTGTAGCACTGTG	CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC

Figure 2. Inhibitory domain mutation is not associated with aberrant joining

(A) Signal joints obtained from extrachromosomal assays with wild-type RAGs. Nonamer and heptamer sequences associated with the 12-RSS and 23-RSS are indicated in boldface type; the number of independent clones represented by each sequence are indicated on the right. Insertions and deletions (dashes) are indicated in red type.

(B) Signal joints from RAG-2(388/405A_{18}) as in (A).

Initial definition of the inhibitory domain

The amino-terminal boundary of the inhibitory domain was defined by extending the alanine scanning mutagenesis into the canonical core region of RAG-2 (Figure 1A). RAG-2(370/387A_{18}) exhibited a gain-of-function phenotype comparable to that of RAG-2(388/405A_{18}), whereas RAG-2(352/369A_{18}) supported wild-type levels of recombination and RAG-2(334/351A_{18}) was not recombination-competent (Figure 3A). Mutations carboxy-terminal to residue 405 failed to confer a gain-of-function phenotype (Figure 3A). RAG protein accumulation was comparable between the RAG-2(370/387A_{18}) and RAG-2(388/405A_{18}) mutations (Figure 3B).

Disruption of inhibition uncouples recombination of endogenous gene segments from H3K4me3 recognition

We employed a qualitative assay to determine whether disruption of the putative RAG-2 inhibitory domain could bypass the dependence of RAG activity on H3K4me3 recognition (Figure 3C). To do so, we expressed wild-type RAG-2, RAG-2(388/405A_{18}), RAG-2(W453A), or RAG-2(388/405A_{18}, W453A) in a RAG-2-deficient pro-B cell line using a retroviral vector that confers puromycin resistance. At 25 days of selection, DSP2-to-J_H joining was assayed (Liu et al. 2007). Rearrangements were detected in cells transduced with core RAG-2, wild-type RAG-2, or RAG-2(388/405A_{18}) (Figure 3D, lanes 3-5). Rearrangement was profoundly reduced in cells transduced with RAG-2(W453A) (Figure 3D, lane 6). The debilitating effect of the W453A mutation was reversed, however, but the second site mutation of residues 388-405 (Figure 3C, lane 7). Consistent with the interpretation that relief of autoinhibition bypasses the dependence of

endogenous V(D)J recombination on recognition of H3K4me3 by RAG-2. This result was reproducible in multiple southern blots and using a different primer pair (results not shown).

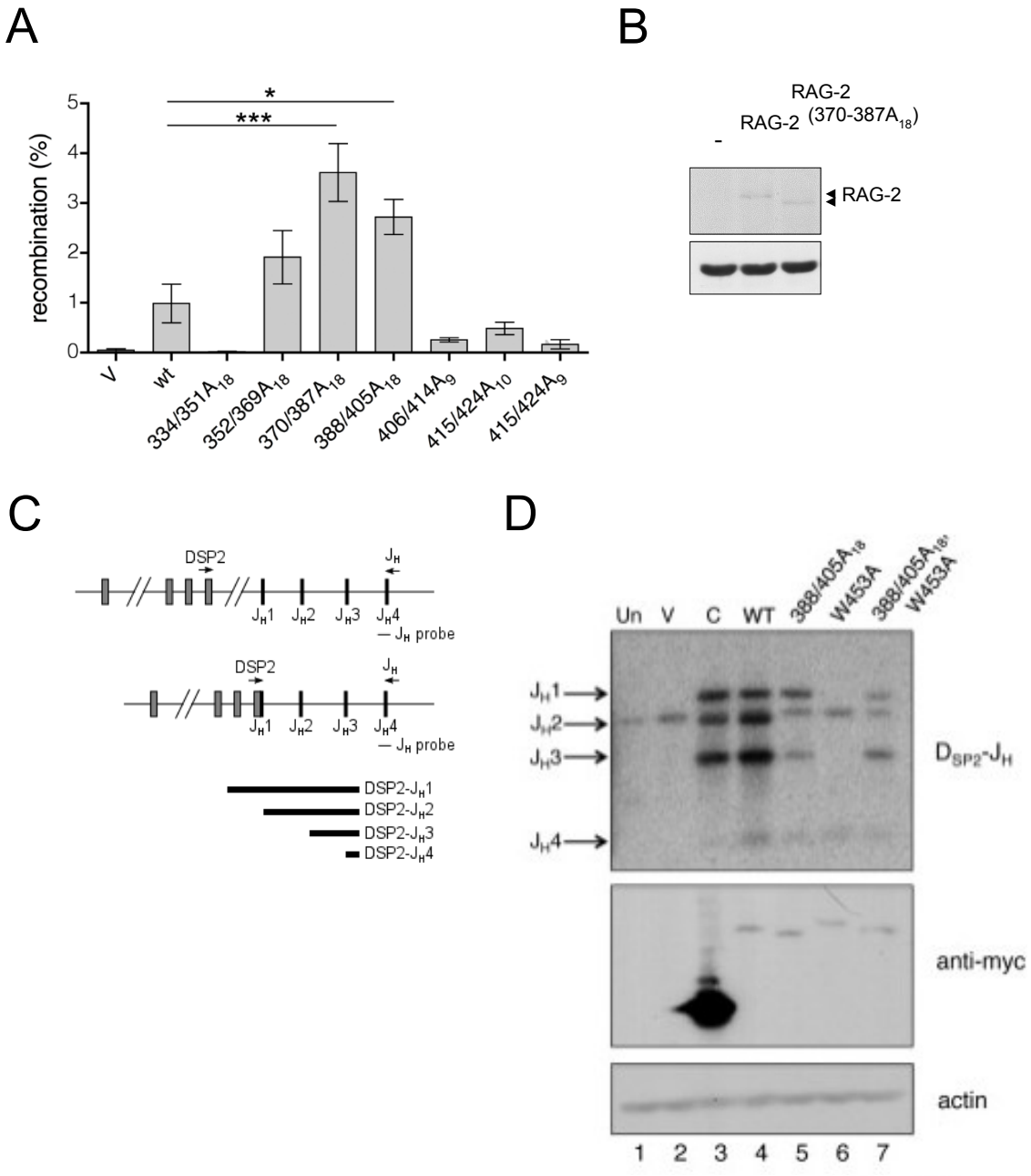

A
recombination (%)
V
wt
334/351A18
352/369A18
370/387A18
388/405A18
406/414A9
415/424A10
415/424A9
B
RAG-2
RAG-2 (370-387A18)
RAG-2
C
DSP2
JH
JH1 JH2 JH3 JH4
JH probe
DSP2-JH1
DSP2-JH2
DSP2-JH3
DSP2-JH4
D
Un V C WT 388/405A18 W453A 388/405A18 W453A
DSP2-JH
anti-myc
actin
1 2 3 4 5 6 7

Figure 3. Initial localization of the RAG-2 inhibitory domain

(A) Mapping of the inhibitory domain. Signal joining (mean ± SEM, n = 3 independent biological replicates, representing ≥ 100 A-resistant colonies per RAG-2 variant) was assayed using full-length RAG constructs and analyzed by ANOVA, ***$p < 0.001$; *$p < 0.05$.

(B) Immunoblotting of RAG-2(370/387A_{18}) and wild-type RAG-2 in transfected NIH3T3 cells. RAG-2 (upper panel) and actin (lower panel) were detected by anti-myc and anti-actin respectively.

(C) Representation of a portion of the IgH locus with relative positions of primers and the probe used to assay endogenous DSP2-J_H joining.

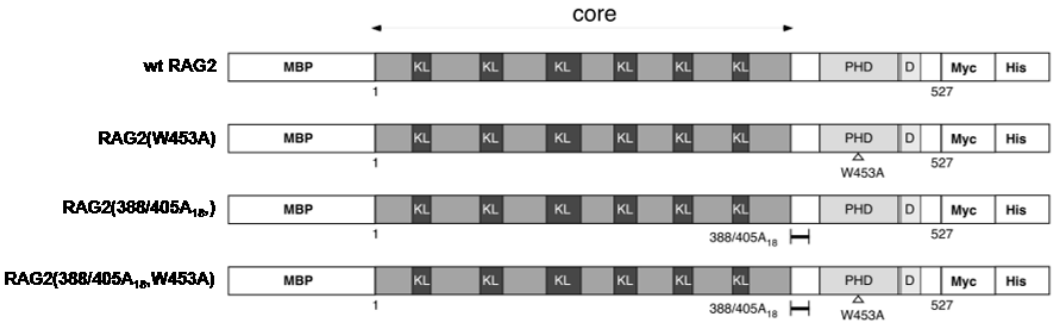
(D) Rescue of W453A by 388/405A_{18} in an assay for endogenous rearrangement. Top: assay for DSP2-J_H joints in genomic DNA from uninfected cells (Un) or cells transduced with the following: vector (V), core RAG-2 (C), wild-type full-length RAG-2 (WT), or full-length RAG-2 mutants. Positions of the DSP2-J_H recombinants are indicated at left. Middle: immunodetection of myc-tagged RAG-2 species with an anti-myc antibody. Bottom: Immunoblotting of actin as a loading control.

The RAG-2 inhibitory domain mutant exhibits increased basal activity, but remains responsive to H3K4me3 (in collaboration with Chao Lu)

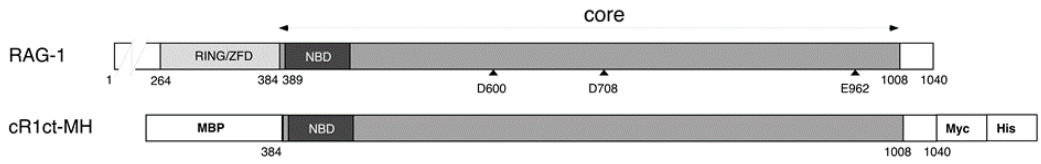
The ability of the 388/405A_{18} mutation to rescue activity of a PHD finger mutant was consistent with (1) disruption of an inhibitory domain whose action in the wild-type protein is relieved by H3K4me3 or (2) disruption of a separate mode of autoinhibition whose action is independent of H3K4me3 binding. To test these possibilities, we assayed wild-type RAG-2, RAG-2(W453A), RAG-2(388/405A_{18}), and RAG-2(388/405A_{18}, W453A) (Figure 4A) with c RAG-1ct (Figure 4B) for responsiveness to H3K4me3 in a coupled cleavage assay. Equivalent amounts of active tetramer, as determined by burst kinetic analysis (Figure 5) were assayed for coupled cleavage of a radiolabeled 12-RSS in the presence of increasing concentrations of H3K4me0 or H3K4me3 (Figure C). As expected, H3K4me3 stimulated hairpin formation by wild-type RAG: at 4 μM H3K4me3, the yield of hairpin product was more than 10-fold greater than in the absence of the peptide (Figure 4D, right). Stimulation was specific to H3K4me3, as H3K4me0 had no effect on coupled cleavage activity (Figure 4D, left). RAG-2(W453A) exhibited basal activity comparable to that of wild-type RAG-2 (Figure 4D, right), indicating that an intact PHD finger is required for stimulation. Consistent with its ability to rescue the function of RAG-2(W453A) *in vivo*, the 388/405A_{18} mutation was associated with increased basal cleavage activity, either alone or in combination with W453A (Figure 4D, left). Despite this increase in basal activity, RAG-2(388/405A_{18}) could be stimulated by H3K4me3 (Figure 4D); responsiveness required an intact PHD finger, as RAG-2(388/405A_{18}, W453A) was not stimulated (Figure 4D, right). Consistent with these observations, RAG-2 fragments bearing the W453A and the 388/405A18, W453A double

mutation failed to bind H3K4me3, whereas a fragment bearing the 388/405A18 mutation retained the ability to bind (Figure 6A, B). Altogether, these observations indicate that the 388/405A18 gain-of-function mutation confers increased basal cleavage activity *in vitro*, but this mutation spares one or more additional inhibitory functions that can be relieved by H3K4me3 binding.

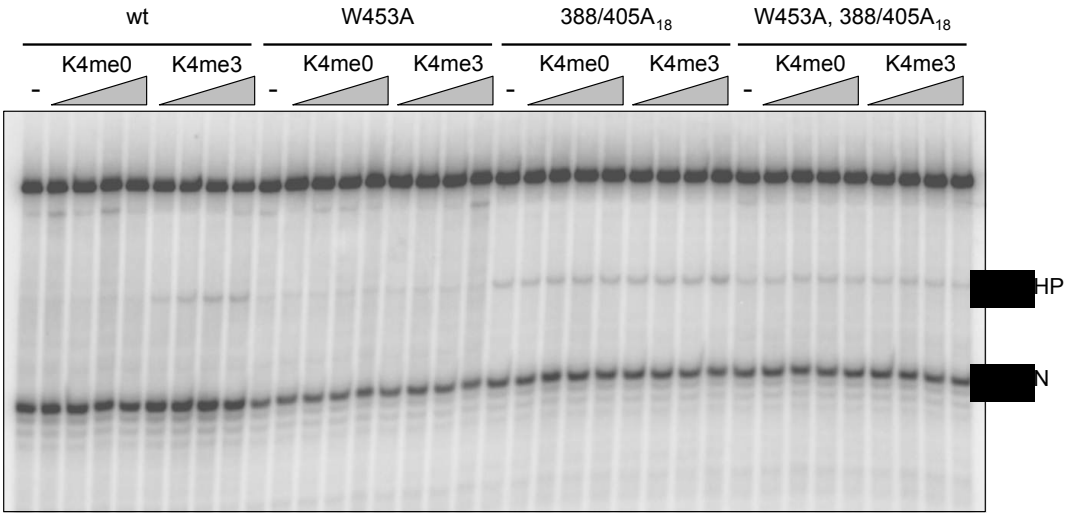
A


core
wt RAG2
MBP
KL
PHD
D
Myc
His
1
527
RAG2(W453A)
W453A
RAG2(388/405A18)
388/405A18
RAG2(388/405A18,W453A)

B


core
RAG-1
RING/ZFD
NBD
1
264
384
389
D600
D708
E962
1008
1040
cR1ct-MH
MBP
Myc
His

C


wt
W453A
388/405A18
W453A, 388/405A18
K4me0
K4me3
-
HP
N

D

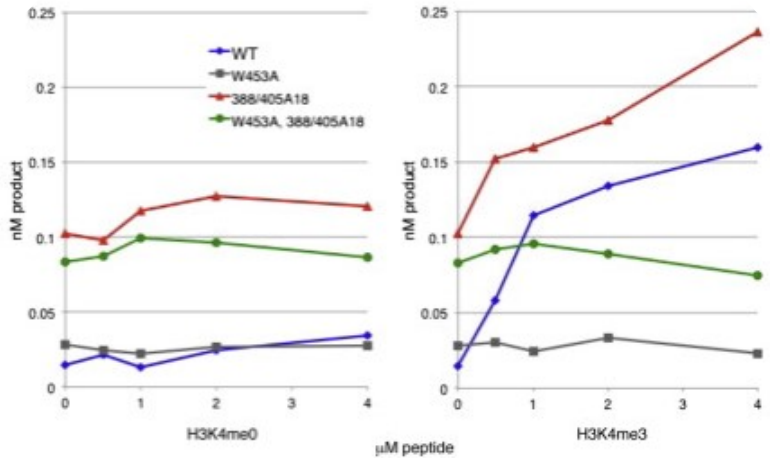

WT
W453A
388/405A18
W453A, 388/405A18
nM product
0.25
0.2
0.15
0.1
0.05
0
0
1
2
3
4
H3K4me0
H3K4me3
μM peptide

Figure 4. Mutation of the RAG-2 inhibitory domain increases coupled cleavage *in vitro*

(A) Diagrams of RAG-2 constructs. Amino acid residues at domain boundaries are numbered. The core is shown in dark gray with Kelch-like domains (KL) indicated. The PHD finger (PHD) and degradation signal (D) in the noncore region are indicated. MBP, Myc, and His denote the maltose binding protein, c-*myc* epitope, and polyhistidine tags, respectively.

(B) Diagrams of wild-type RAG-1 (top) and core RAG-1ct-MH (cR1ct-MH, bottom). Amino acid residues at domain boundaries are numbered. The core is designated in dark gray. The RING-type zinc-finger domain (RING/ZFD) and the nonamer binding domain (NBD) are indicated. Arrowheads mark the catalytic residues.

(C) Coupled cleavage reactions contained radiolabeled 12-RSS, unlabeled 23-RSS, and wild-type (WT) RAG-2 or RAG-2 mutants as defined at top. K4me0 and K4me3, reactions supplemented with 0.5, 1, 2, or 4 μM H3K4me0 or H3K4me3 peptide; −, reactions lacking peptide. Positions of hairpin (HP) and nicked (N) products are indicated by arrows.

(D) Accumulation of hairpin product at 1 hr (nM product) is plotted as a function of the concentration of H3K4me0 (left) or H3K4me3 (right). Blue diamonds, wild-type RAG-2; gray squares, RAG-2(W453A); red triangles, RAG-2(388/405A_{18}); green circles, RAG-2(W453A, 388/405A_{18}).

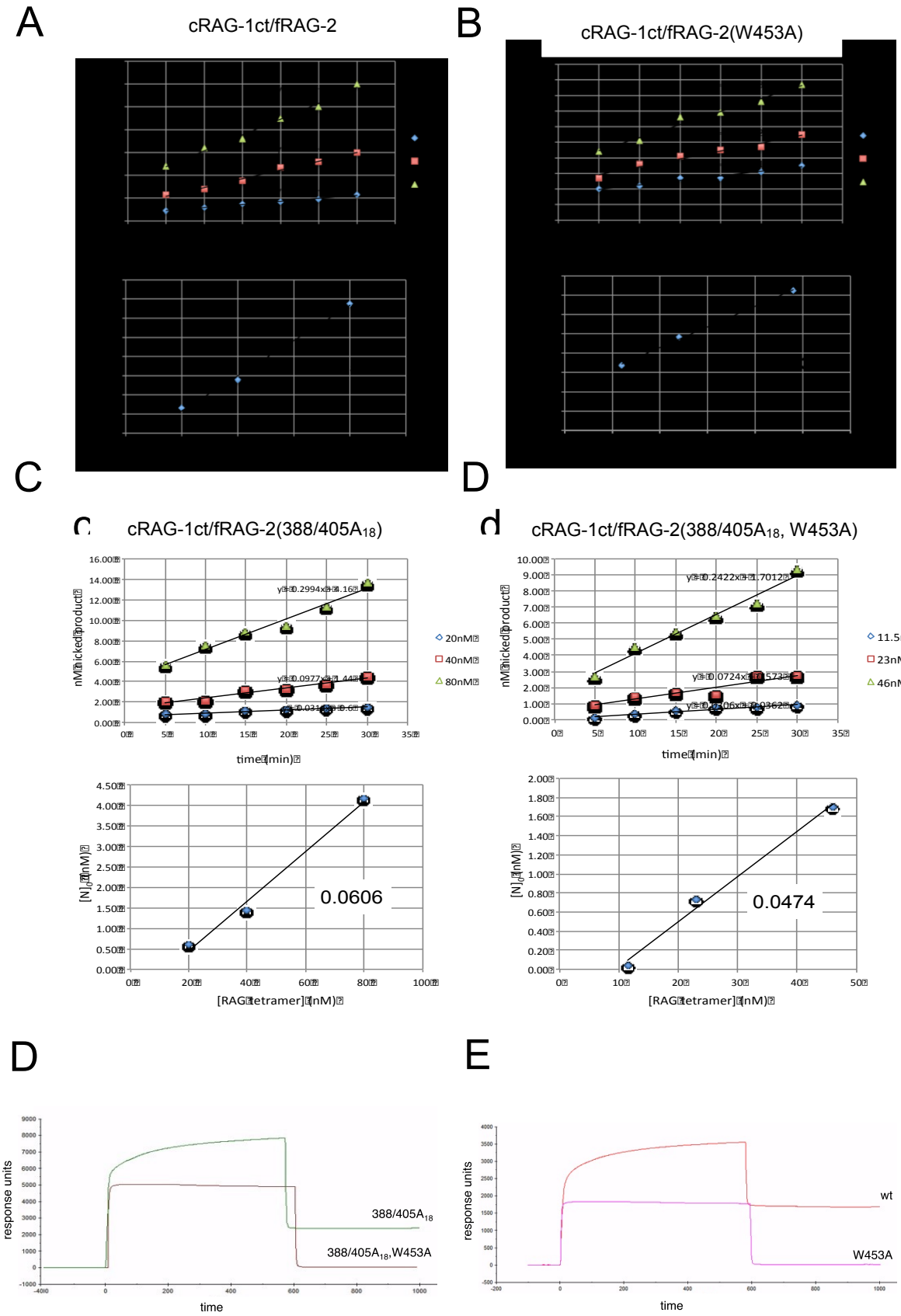

A
cRAG-1ct/fRAG-2
B
cRAG-1ct/fRAG-2(W453A)
C
cRAG-1ct/fRAG-2(388/405A18)
D
cRAG-1ct/fRAG-2(388/405A18, W453A)
time (min)
20nM
40nM
80nM
11.5nM
23nM
46nM
0.0606
0.0474
[RAG tetramer] (nM)
D
E
response units
time
388/405A18
388/405A18,W453A
wt
W453A

Figure 5. Burst kinetic and surface plasmon resonance analyses of wild-type and mutant RAG complexes

(A) Complexes of cRAG-1ct with full-length wild-type RAG-2 were assayed for nicking of a 12-RSS substrate at various nominal RAG concentrations. Upper panels, accumulation of nicked product as a function of time; nominal RAG concentration, calculated based on tetrameric stoichiometry, is indicated at right. Lower panels, estimation of active fraction. The kinetic curves of the upper panels were extrapolated to zero time to give the $[N]_0$ associated with each nominal RAG concentration ([RAG tetramer]). $[N]_0$ was then plotted as a function of [RAG tetramer]; the slope of each resulting curve represents the fraction of active RAG in the preparation.

(B) RAG-2(W453A) analyzed as in (A).

(C) RAG-2(388/405A$_{18}$) analyzed as in (A).

(D) RAG-2(388/405A$_{18}$, W453A) analyzed as in (A). The active fractions of the four preparations were similar, ranging from 4.6 percent to 6.1 percent.

(E) Surface plasmon resonance assays for the association of RAG-2PHD(388/405A$_{18}$) and RAG-2PHD(388/405A$_{18}$, W453A) with a histone H3K4me3 peptide. Biotin-tagged H3K4me3 peptide (residues 1 – 21) was immobilized at 24 pmol on biosensor chips. GST-tagged RAG-2PHD variants were brought to 2 μM and injected for 10 min at a flow rate of 5 μl/min. Subsequently, running buffer alone was injected for 10 min.

(F) Surface plasmon resonance as in (E) for wild-type RAG-2PHD and RAG-2PHD(W453A).

H3K4me3 and inhibitory domain mutation stimulate substrate binding (in collaboration with Chao Lu)

The stimulatory effect of the 388/405A_{18} mutation could result from increased affinity for substrate, increased catalytic activity, or both. To distinguish these possibilities, we assessed substrate binding and catalysis. To measure affinity for DNA substrate, a 12-RSS fragment was incubated with increasing concentrations of wild-type RAG in the presence of 4 μM H3K4me0 or H3K4me3 peptide. Incubation was carried out in the presence of Ca^{2+}, which supports the binding of RAG to substrate in the absence of DNA cleavage. The fraction of total substrate remaining in the unbound state was determined (Figure 6A) and expressed as a function of active RAG concentration (Figure 6C), as defined by burst kinetics under the assumption that the active unit is a heterotetramer of composition (RAG-1)2(RAG-2)2 (Yu & Lieber 2000). Dissociation constants, K_D, were determined (see the Materials and Methods). The addition of H3K4me3 was accompanied by an increase in the affinity of wild-type RAG for substrate DNA, relative to H3K4me0; in contrast, the affinity of RAG-2(388/405A18) for a 12-RSS substrate was similar in the presence of H3K4me0 or H3K4me3 (Figures 6B and 3D). We also performed direct comparisons of substrate binding by each RAG species in the presence of H3K4me0 (Figure 6E) or H3K4me3 (Figure 6F). In the presence of control peptide, the affinities (K_D) of RAG-2(388/405A18) and wild-type RAG for substrate were estimated at 88 nM and 242 nM, respectively (Figure 6G). In the presence of H3K4me3, wild-type RAG and RAG-2(388/405A18) bound substrate with an estimated KD of 76 nM and 84 nM, respectively (Figure 6H), similar to the affinity of RAG-2(388/405A18) for substrate in the presence of H3K4me0. Thus the 388/405A18

mutation confers a constitutive increase in substrate binding affinity by RAG independent of the presence of H3K4me3.

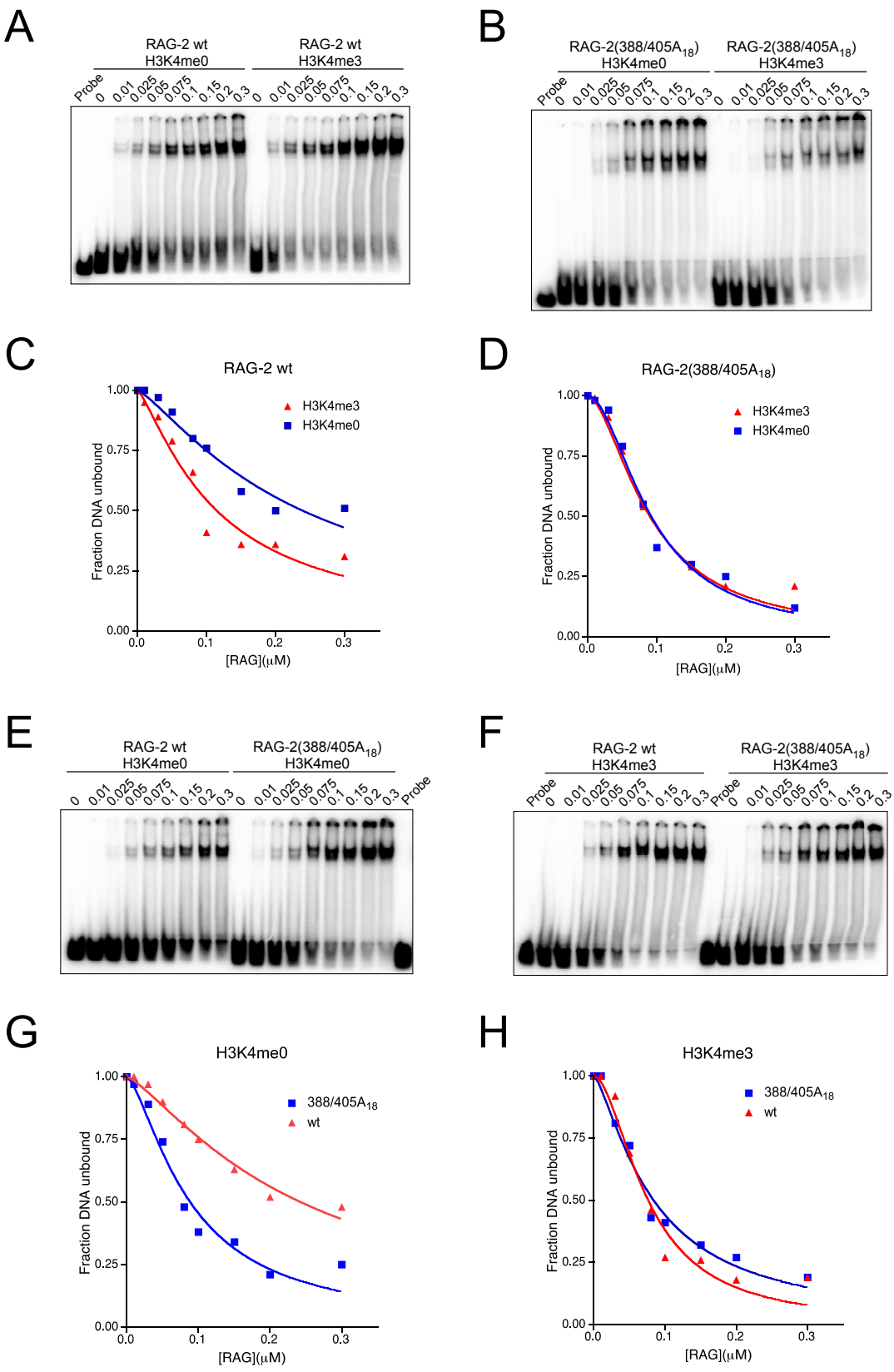

A
RAG-2 wt
H3K4me0
RAG-2 wt
H3K4me3
Probe
0 0.01 0.025 0.05 0.075 0.1 0.15 0.2 0.3
B
RAG-2(388/405A18)
H3K4me0
RAG-2(388/405A18)
H3K4me3
C
RAG-2 wt
H3K4me3
H3K4me0
Fraction DNA unbound
[RAG](μM)
D
RAG-2(388/405A18)
E
RAG-2 wt
H3K4me0
RAG-2(388/405A18)
H3K4me0
F
RAG-2 wt
H3K4me3
RAG-2(388/405A18)
H3K4me3
G
H3K4me0
388/405A18
wt
H
H3K4me3

Figure 6. The RAG-2(388/405A$_{18}$) mutation stimulates RAG-RSS binding

(A) Electrophoretic mobility shift assay (EMSA) for binding of wild-type RAG to a consensus 12-RSS in the presence of H3K4me0 or H3K4me3 peptide as indicated at top. Probe shows 12 RSS incubated in the absence of RAG. The concentration, in μM, of active RAG in each binding reaction is indicated above the lane.

(B) EMSA as in (A) except that RAG-2(388/405A18) was substituted for wild-type RAG-2.

(C) H3K4me3 reduces the K_D of RAG-RSS binding. The fraction of free probe (fraction DNA unbound) in each binding reaction of (A) was plotted as a function of active RAG concentration. Data from reactions containing H3K4me0 and H3K4me3 are indicated by blue squares and red triangles, respectively.

(D) The RAG-2 388/405A18 mutation relieves responsiveness of RAG-RSS binding to H3K4me3. The fraction of free probe (fraction DNA unbound) in each binding reaction of (B) is plotted as in (C).

(E) EMSA for binding of wild-type RAG (left) or RAG-2(388/405A18) (right) to a consensus 12-RSS in the presence of H3K4me0. Probe, 12 RSS incubated in the absence of RAG. The concentration of active RAG in each reaction is indicated above the lane.

(F) EMSA as in (E), except that H3K4me3 was substituted for H3K4me0.

(G) The RAG-2 388/405A18 mutation increases basal affinity of RAG for RSS in the absence of H3K4me3. The fraction of free probe (fraction DNA unbound) in each binding reaction of (E) was plotted as a function of active RAG concentration. Data from reactions containing RAG-2(388/405A18) and wild-type RAG-2 are indicated by blue squares and red triangles, respectively.

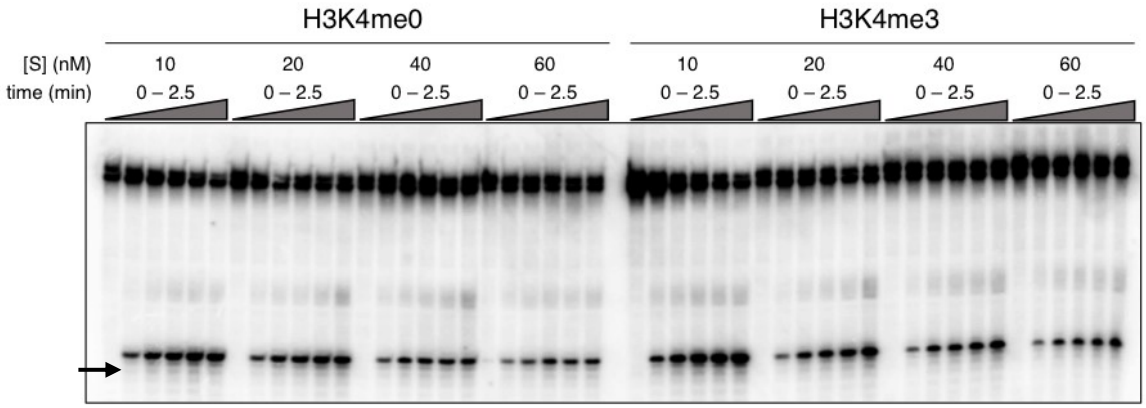
(H) The fraction of free probe (fraction DNA unbound) in each binding reaction of (F) is plotted as in (G).

Data in (A), (B), (E), and (F) are representative of three experiments.

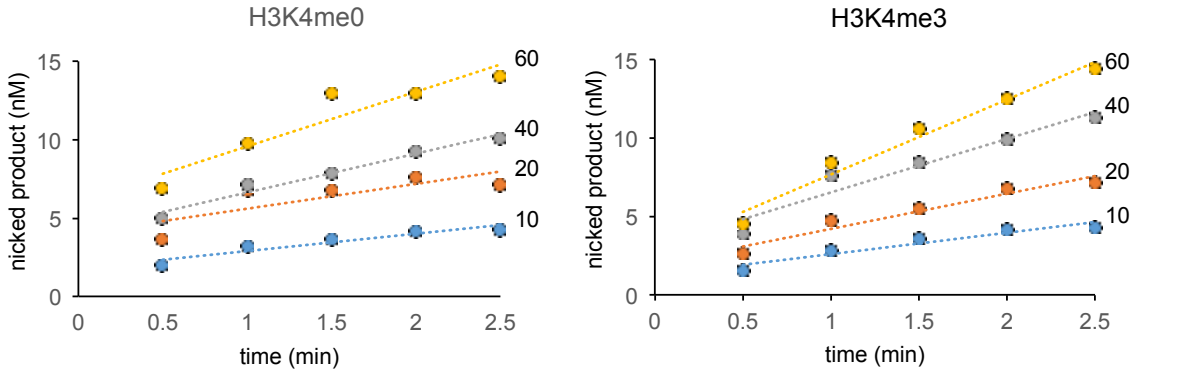
H3K4me3 and the 388/405A$_{18}$ mutation stimulate RAG catalysis (in collaboration with Chao Lu)

Because the 388/405A$_{18}$ mutation uncoupled the high affinity state from H3K4me3 binding, we could assess the effect of H3K4me3 on k_{cat} in the absence of its effect on K_D. We assayed nicking of a pre-bound 12-RSS substrate at concentrations from 10 nM to 60 nM by RAG-2(388/405A$_{18}$) in complex with cR1ct-MH. Reactions were carried out at an active RAG tetramer concentration of 1.5 nM in the presence of 4 μM H3K4me0 or H3K4me3 peptide (Figure 7A). Following determination of V_{max} (Figures 7B and 7C), k_{cat} was estimated (Materials and Methods). In the presence of H3K4me0, RAG-2(388/405A18) supported nicking with an apparent k_{cat} of 4.95 min^{-1}, which increased to 7.06 min^{-1} in the presence of H3K4me3 (Figure 7C, D). In comparison, we observed turnover rates of 0.83 min−1 and 3.76 min−1 for wild-type RAG-2 in the presence of H3K4me0 or H3K4me3, respectively (Figure 7D), consistent with previous estimates (Shimazaki et al. 2009). Thus, the 388/405A$_{18}$ mutation is associated not only with increased affinity for substrate but also with a 6-fold increase in the basal k_{cat} for DNA nicking; the basal k_{cat} observed for RAG-2(388/405A$_{18}$) is similar to that observed for wild-type RAG-2 in the presence of H3K4me3. Nonetheless, RAG-2(388/405A$_{18}$) remains able to respond to H3K4me3 with an increase in catalytic rate.

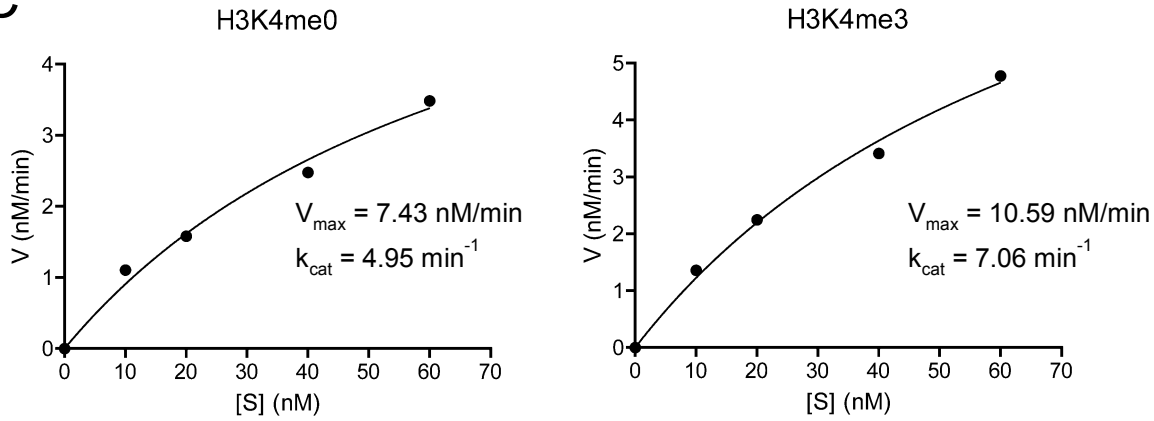
A



B



C



D

	RAG-2		RAG-2(388/405A$_{18}$)	
	+H3K4me0	**+H3K4me3**	**+H3K4me0**	**+H3K4me3**
K_D	242 nM	76 nM	88 nM	84 nM
K_{cat}	0.83 min^{-1}	3.76 min^{-1}	4.95 min^{-1}	7.06 min^{-1}

Figure 7. Disruption of the RAG-2 inhibitory domain and H3K4me3 binding both stimulate RAG catalysis

(A) Assay for RSS nicking. Reactions contained 1.5 nM RAG-2(388/405A_{18}) and 12-RSS substrate HL44/45 at 10, 20, 40, or 60 nM. Reactions were supplemented with 4 μM H3K4me0 or H3K4me3 peptide as indicated at top. Accumulation of nicked product (arrow) was assayed at times ranging from 0 to 2.5 min.

(B) Concentration of nicked product as determined in (A) is plotted against time for each substrate concentration. Blue, 10 nM; orange, 20 nM; gray, 40 nM; and yellow, 60 nM. Left: reactions containing H3K4me0; right: reactions containing H3K4me3.

(C) Reaction velocity (V) is plotted in nM/min as a function of substrate concentration ([S]). V_{max} was determined by nonlinear regression analysis (Materials and Methods); $k_{cat} = V_{max}/[RAG]_T$, where $[RAG]_T$ is the total concentration of active RAG tetramer.

(D) Estimates of dissociation constants (K_D) and catalytic rate (k_{cat}) for full-length wild-type RAG-2 and full-length RAG-2(388/405A_{18}) in complex with cR1ct-MH. K_D was measured for a canonical 12-RSS substrate; k_{cat} was determined in an assay for nicking of a 12-RSS substrate.

Characterization of the inhibitory domain within the noncore region of RAG-2

The RAG-2 W453A mutation disrupts the PHD finger, abolishes specific binding of the recombinase to H3K4me3 and impairs recombination *in vivo*. Contiguous mutations of RAG-2 from amino acid 370 to 405 stimulated the activity of RAG-2(W453A) (Figure 3A). This interval is predominantly acidic with interspersed proline, serine, threonine and phenylalanine residues (Figure 8A). We wished to determine the relative contributions of these residues to inhibition of RAG recombination activity. This was of particular interest because structural analysis suggested that the RAG-2 PHD finger could engage a proline residue in the absence of H3K4me3 (Ramón-Maiques et al. 2007). Mutation of the proline residues resulted in no difference in recombination activity (Figure 8B). Mutation of the serine and threonine or phenylalanine (Figure 8C) residues conferred only modest increases in recombination activity. In contrast, neutralization of acidic residues in the 370-405 interval RAG-2(D/E370-405A) increased recombination activity to the level observed for RAG-2(370-387A_{18}) (Figure 8C). These differences in recombination activity were not explained by differences in protein expression (Figure 8D, E).

A

RAG2 352-EEDLSEDQKIVSNSQTSTEDPGDSTPFEDSEEFCFSAEATSFDGDDEFDTYNED-405

RAG2(387-405A$_{18}$) EEDLSEDQKIVSNSQTSTEDPGDSTPFEDSEEFCFS**AAAAAAAAAAAAAAAAAA**

RAG2(P372,377A) EEDLSEDQKIVSNSQTSTED**A**GDST**A**FEDSEEFCFSAEATSFDGDDEFDTYNED

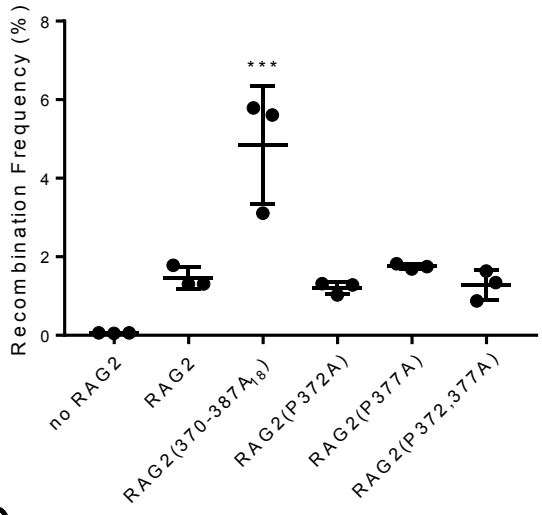
RAG2(F370-405A) EEDLSEDQKIVSNSQTSTEDPGDSTP**A**EDSEE**A**C**A**SAEATS**A**DGDDE**A**DTYNED

RAG2(S/T370-405A) EEDLSEDQKIVSNSQTSTEDPGD**AA**PFED**A**EEFCF**A**AEA**AA**FDGDDEFD**A**YNED

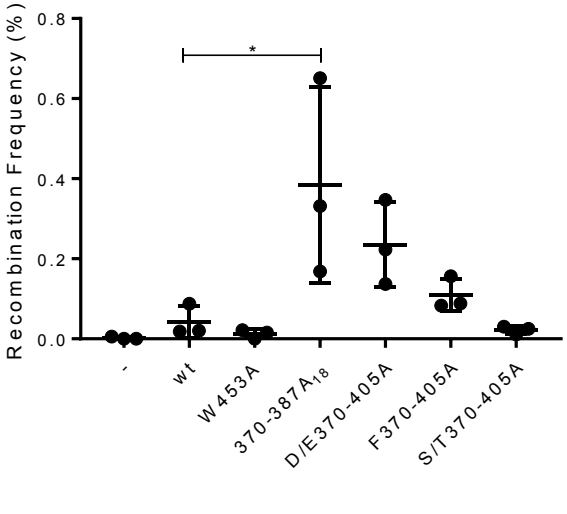
RAG2(D/E370-405A) EEDLSEDQKIVSNSQTST**AA**PG**A**STPF**AA**S**AA**FCFSA**A**ATSF**A**G**AAA**F**A**TYN**AA**

RAG2(D/E352-405A) **AAA**LS**AA**QKIVSNSQTST**AA**PG**A**STPF**AA**S**AA**FCFSA**A**ATSF**A**G**AAA**F**A**TYN**AA**

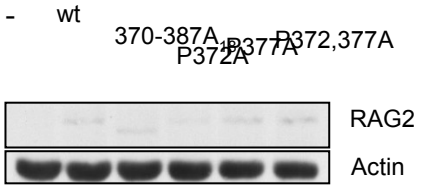
B



C



D



E

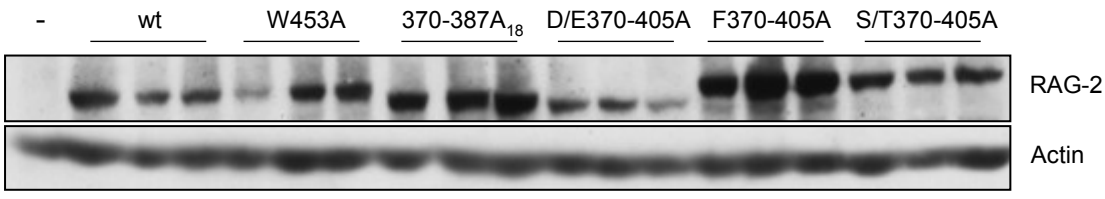


Figure 8. Localization of the RAG-2 inhibitory domain to the acidic residues spanning amino acid 352 to 405

(A) Diagram of RAG-2 with the canonical core (CORE) and plant homeodomain (PHD) indicated. Amino acids at domain boundaries are numbered. The location of W453 is noted with a dotted line. T490 is noted with a solid line. Below: Wild-type RAG-2 sequence from amino acid 352 through 405 (top) compared with the sequences of RAG-2 mutants. Mutated residues are indicated in bold type.

(B) Activity of RAG-2 proline mutations on an extrachromosomal signal joint substrate, pJH200. Recombination frequencies from individual transfections are plotted with mean ± standard deviation (SD). Brown-Forsythe test was used to ensure SDs were not significantly different ($p > 0.05$). One-way ANOVA was used with Tukey's multiple comparison post-test. *** sample was significantly different ($p > 0.001$) from all other samples and was the only significant difference.

(C) Activity of RAG-2 mutations on an extrachromosomal substrate as in (B). Not all significant differences are indicated. * notes that means that were significantly different ($p > 0.05$).

(D) Top panel: Immunodetection of RAG-2 variants in (B) using an anti-myc antibody. Bottom panel: Actin, detected with anti-actin, is used as a loading control.

(E) Immunodetection as in (D) for the RAG-2 variants in (C).

The RAG-2 inhibitory domain acts independently of the PHD

The predominance of acidic residues extends to residue 352, therefore we asked whether inhibition of RAG activity could be further relieved by truncation into this region (Figure 9A). Carboxy-terminal truncation of RAG-2 from residue 405 to 351 was associated with a significant ($p < 0.01$) increase in recombination (Figure 9B) activity despite decreased protein expression (Figure 9E). Moreover, recombination activity was increased when neutralization of acidic residues was extended from residue 370 to 352 (ns, $p = 0.0891$) (Figure 9C). Prior work suggested that further neutralization would impair recombination activity, as RAG-2(334-351A_{18}) was inactive on an extrachromosomal substrate (Figure 3A). In all subsequent experiments to test relief of inhibition, we used the optimally active RAG variant, RAG-2(D/E352-405A).

Importantly, the stimulatory effects of truncation on recombination suggested that the inhibitory domain does not exert its suppressive effects solely through an interaction with the PHD finger (Figure 9B). To confirm this, we mutated the inhibitory domain in the context of a RAG-2 construct lacking the PHD finger. The resulting mutant, RAG-2(1-405, D/E352-405A), exhibited a significant increase in function relative to the RAG-2(1-405) mutant (Figure 9D). None of these differences in activity could be attributed to changes in protein expression (Figure 9F, G).

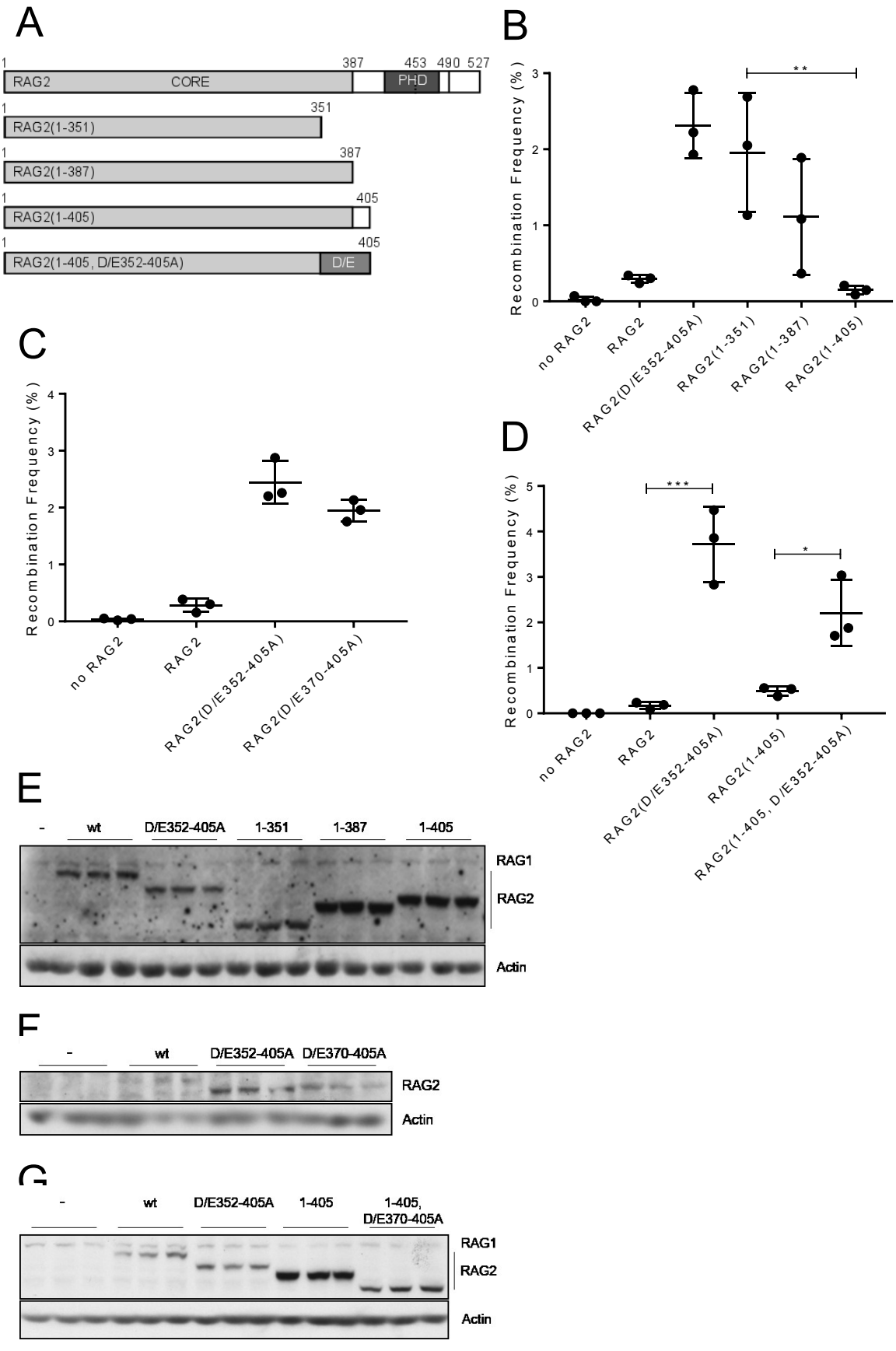

A
1
387
453
490
527
RAG2
CORE
PHD
1
351
RAG2(1-351)
1
387
RAG2(1-387)
1
405
RAG2(1-405)
1
405
RAG2(1-405, D/E352-405A)
D/E
B
Recombination Frequency (%)
**
no RAG2
RAG2
RAG2(D/E352-405A)
RAG2(1-351)
RAG2(1-387)
RAG2(1-405)
C
Recombination Frequency (%)
no RAG2
RAG2
RAG2(D/E352-405A)
RAG2(D/E370-405A)
D
Recombination Frequency (%)
***
*
no RAG2
RAG2
RAG2(D/E352-405A)
RAG2(1-405)
RAG2(1-405, D/E352-405A)
E
-
wt
D/E352-405A
1-351
1-387
1-405
RAG1
RAG2
Actin
F
-
wt
D/E352-405A
D/E370-405A
RAG2
Actin
G
-
wt
D/E352-405A
1-405
1-405, D/E370-405A
RAG1
RAG2
Actin

Figure 9. The RAG-2 inhibitory domain acts independent of the PHD

(A) Diagram of the RAG-2 truncations used in this study. Numbers indicate the boundaries of defined domains. The canonical core (CORE) is in light gray with the plant homeodomain (PHD) in dark gray. The position of W453 is marked with a dashed line and the location of T490 is marked with a solid line. D/E in medium gray indicates a truncation at 405 combined with the D/E352-405A inhibitory domain mutation.

(B) The inhibitory domain spanning residues 352 through 405 functions in the absence of the PHD. Recombination frequency plotted for three independent transfections of the pJH200 extrachromosomal substrate, RAG-1, and the indicated RAG-2 variant. Brown-Forsythe test was used to ensure SDs were not significantly different ($p > 0.05$). One-way ANOVA was used with Tukey's multiple comparison post-test. All significant differences not shown, ** indicates the means are significantly different ($p < 0.01$).

(C) Comparison of RAG-2 inhibitory domain neutralization mutations. Analyzed as in (B). Both mutations were significantly higher than wild-type RAG-2 ($p < 0.0001$), but not significantly different from each other ($p = 0.0891$).

(D) Neutralization of the interval between 352 and 405 relieves inhibition in the context of truncated RAG-2. Analyzed as in (B). All significant differences not shown, * indicates means are significantly different ($p < 0.05$), *** indicates means are significantly different ($p< 0.001$).

(E), (F), and (G) represent immunodetection of RAG-2 variants for (B), (C), and (D) respectively with an anti-myc antibody (top) and actin as a loading control (bottom).

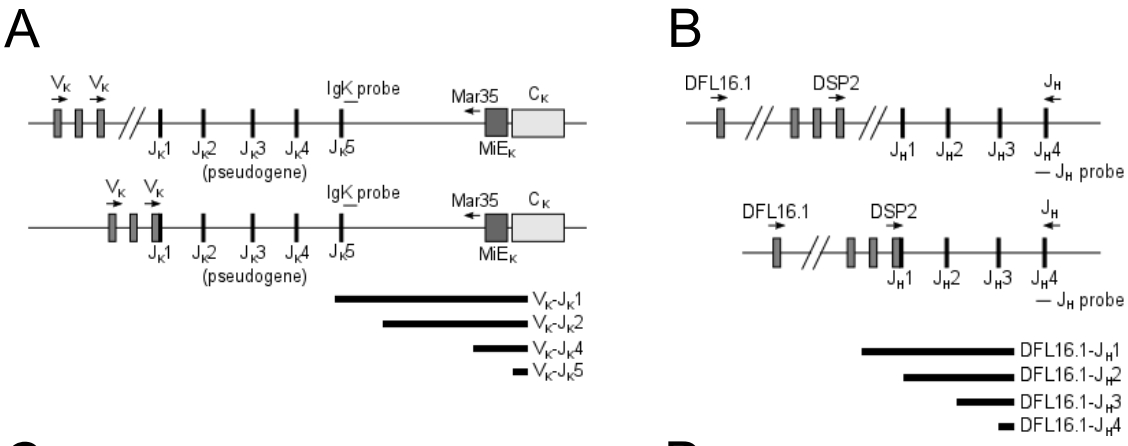
The RAG-2 inhibitory domain suppresses recombination in B progenitor cells

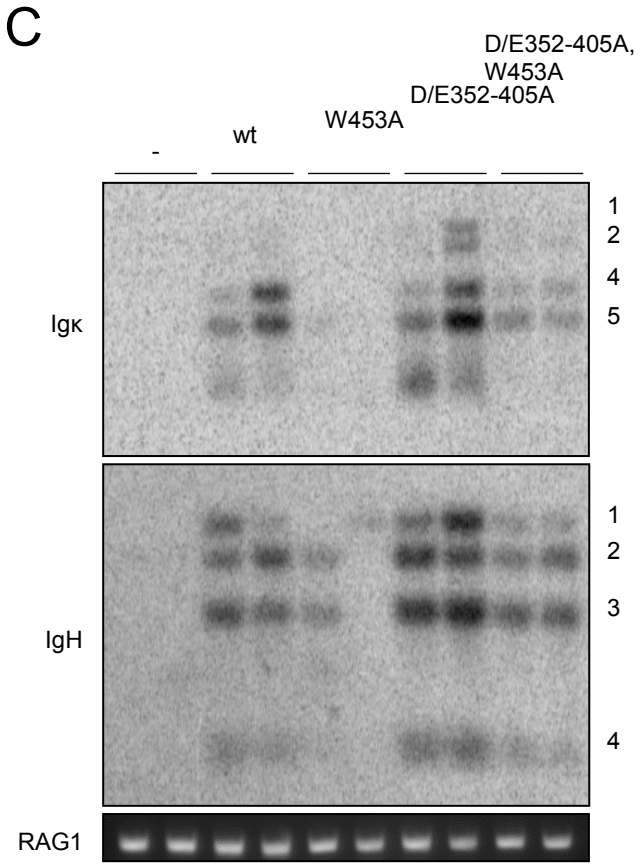
We asked whether disruption of this inhibitory domain would affect rearrangement of endogenous immunoglobulin loci. Our initial experiments employed the RAG-2-deficient pro-B cell line, 63-12. Wild-type RAG-2 or RAG-2 mutants were introduced by retroviral transduction and both V_κ-to-J_κ (Figure 10A) and D_H-to-J_H joining (Figure 10B) were assayed. In cells transduced with wild-type RAG-2, and not those transduced with vector alone, we observed robust V_κ-to-J_κ (Figure 10C top and 10D, compare - and wt) and D_H-to-J_H joining (Figure 10C middle and 10E, compare - and wt). The RAG-2 W453A mutation impaired recombination at both loci (Figure 10C-E, W453A) and this debilitating effect was reversed by secondary mutation of the inhibitory domain (Figure 10C-E, compare W453A and D/E352-405A, W453A). The observed effects of mutation on recombination were not explained by differences in expression of RAG-2 (Figure 10F).

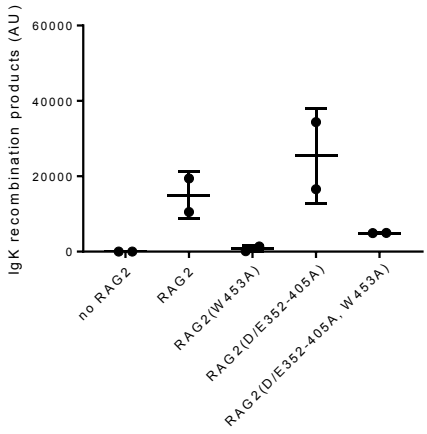
We introduced the same RAG-2 constructs into RAG-2-deficient Abelson murine leukemia virus (AbMuLV)-transformed B progenitor cells in which cell cycle arrest, RAG-1 expression and V(D)J recombination are inducible by treatment with STI-571 (Gapud et al. 2011). In the absence of STI-571 wild-type RAG-2 and RAG-2(D/E352-405A) supported a low level of V_κ-to-J_κ (Figure 11A, B, DMSO lanes) and D_H-to-J_H rearrangement (Figure 11C, D, lanes DMSO lanes) and this increased further after 48 hour treatment with STI-571 (Figure 11A-D, 48hr lanes). Recombination at both loci was impaired by the RAG-2 W453A mutation (Figure 11A-D, lanes W453A), and this impairment was partially reversed by a second-site mutation in the inhibitory domain (Figure 11A-D, lanes D/E352-405A, W453A). The differential effects of these mutations

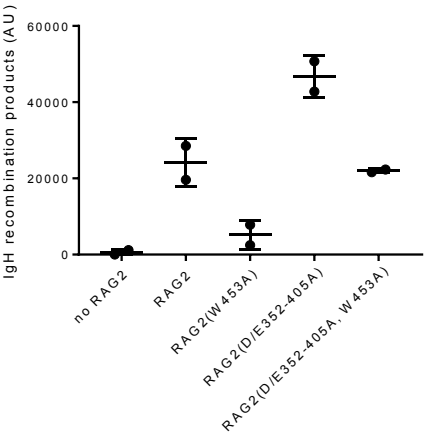
on recombination could not be attributed to differences in expression of RAG-2 protein (Figure 12D) or on differences in germline transcription from the IgH μ0 or Iμ promoters (Figure 11E, F).

The R2K3 pre-B cell line carries an integrated recombination reporter that permits quantitation of recombination (Figure 12A). After introduction of RAG-2 by retroviral transduction and cell cycle arrest with STI-571, inversional joining of recombination signal sequences within the reporter reorients the GFP cassette to permit its expression from an upstream promoter (Figure 12B). As assayed by flow cytometry for fluorescence, no recombination was observed in R2K3 cells lacking RAG-2, either before or after induction (Figure 12B, panel no RAG-2 and Figure 12C). Cells transduced with RAG-2 exhibited a low level of recombination, which increased robustly at 48 and 96 hours after induction by STI-571 (Figure 12B, panel RAG-2 and Figure 12C). Recombination was nearly abolished by the RAG-2 W453A mutation (Figure 12B, panel RAG-2(W453A) and Figure 12C) and this effect was modestly but reproducibly reversed by a second site mutation in the inhibitory domain (Figure 12B, panel RAG-2(D/E352-405A, W453A) and Figure 12C). Mutation of the inhibitory domain alone conferred a reproducible increase in the frequency of recombination relative to wild-type (Figure 12C). These differences were not due to differences in protein expression (Figure 12D) or differences in H3K4me3 levels (Figure 18A).


A
B
IgK probe
Mar35
MiEκ
(pseudogene)
DFL16.1
DSP2
JH probe
DFL16.1-JH1
DFL16.1-JH2
DFL16.1-JH3
DFL16.1-JH4


C
D/E352-405A, W453A
D/E352-405A
W453A
wt
-
Igκ
IgH
RAG1


D
IgK recombination products (AU)
no RAG2
RAG2
RAG2(W453A)
RAG2(D/E352-405A)
RAG2(D/E352-405A, W453A)


E
IgH recombination products (AU)

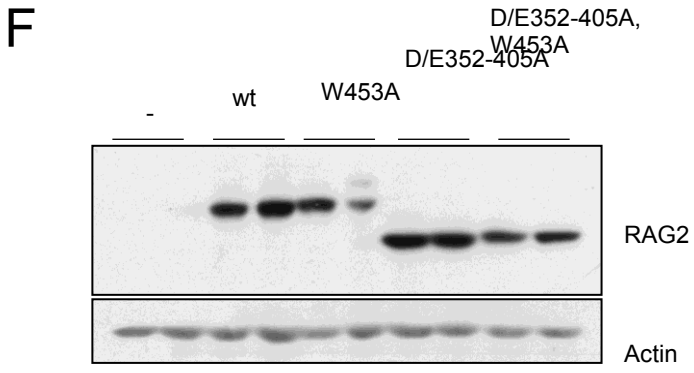

F
D/E352-405A, W453A
D/E352-405A
W453A
wt
-
RAG2
Actin

Figure 10. The RAG-2 inhibitory domain bypasses the requirement for a functional PHD in 63-12 cells

(A) Diagram of the assay for rearrangement of the Igκ locus. A degenerate primer that binds to the seven most commonly used V_k segments was designed (Aoki-Ota et al. 2012) and used with a primer 5' of the κ intronic enhancer (MiEκ) (Inlay et al. 2002). Variable (V), joining (J), and constant (C) regions are indicated. Jκ3 is a pseudogene not used in light chain recombination. After recombination brings the V-primer in proximity of the J-primer, the 4 PCR products are visible by gel electrophoresis and southern blotting.

(B) Diagram of the assay for rearrangement of the IgH locus. DSP2 or DFL16.1 primers were used with a primer 3' of J_H4 to detect D-to-J joining in the IgH locus. As in (A) for the IgH locus.

(C) The RAG-2 inhibitory domain second-site mutation increases recombination activity of RAG-2(W453A) in 63-12 cells. Top: Recombination of the Igκ locus assayed by southern blot, numbers 1 through 4 mark the location of PCR products for the four functional J segments. Middle: Southern blot of endogenous D_H-to-J_H recombination at the IgH locus with the numbers 1 through 4 indicating the positions of products formed with the four functional J segments. Bottom: Ethidium bromide-stained gel showing a PCR of the RAG-1 locus to control for DNA quality.

(D) Quantification of the southern blot for Igκ rearrangement.

(E) Quantification of the southern blot for IgH rearrangement.

(F) Anti-myc immunodetection of myc-tagged RAG-2 (top) and the loading control, actin (bottom).

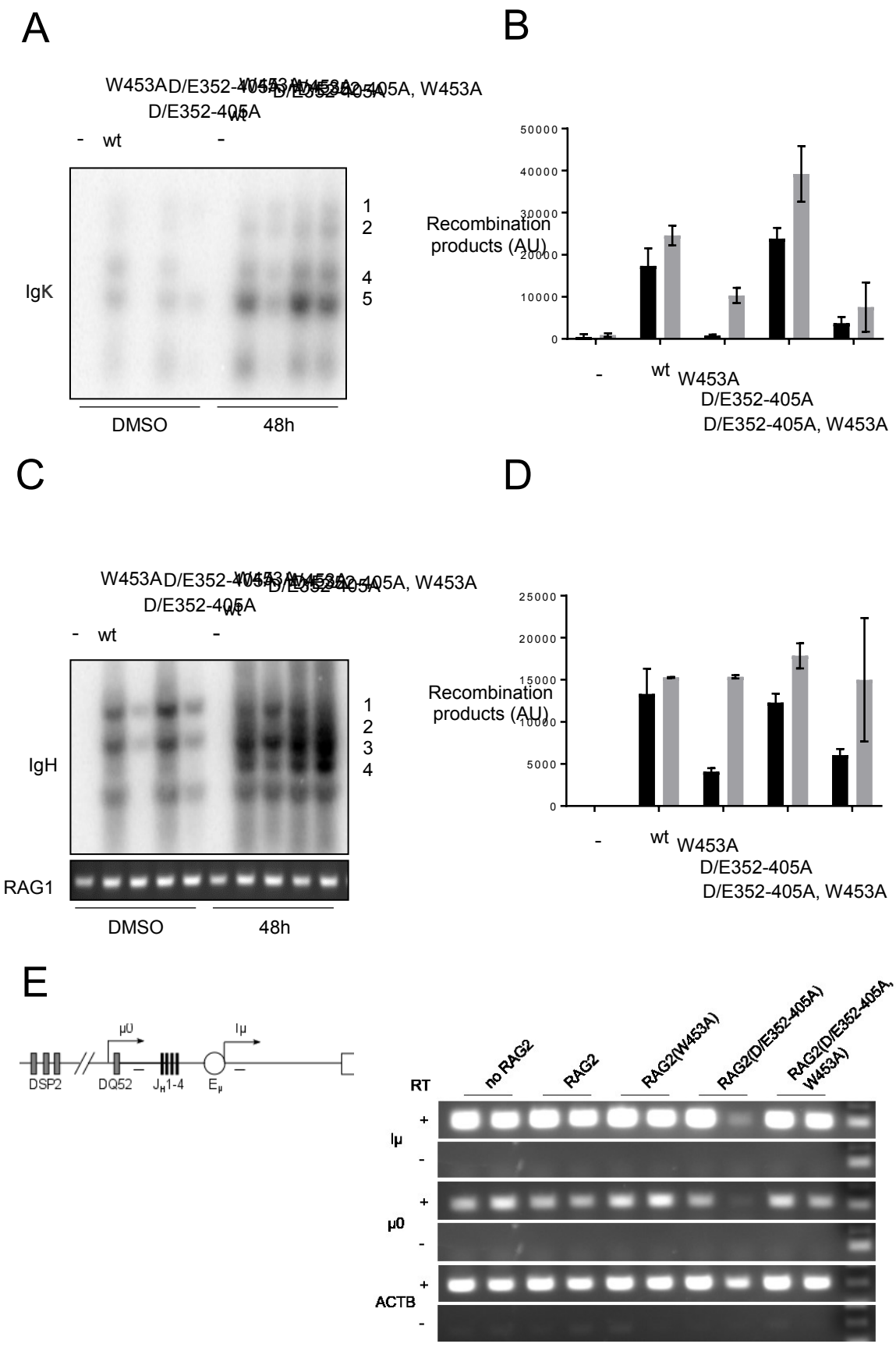

A
B
W453A
D/E352-405A
D/E352-405A, W453A
wt
-
IgK
1
2
4
5
DMSO
48h
Recombination products (AU)
50000
40000
30000
20000
10000
0
C
D
IgH
RAG1
3
25000
20000
15000
10000
5000
E
μ0
Iμ
DSP2
DQ52
J_H1-4
E_μ
RT
no RAG2
RAG2
RAG2(W453A)
RAG2(D/E352-405A)
RAG2(D/E352-405A, W453A)
+
ACTB

Figure 11. The RAG-2 inhibitory domain bypasses the requirement for a functional PHD in R2K3 cells

(A) Igκ recombination in R2K3 cells. R2K3 cells stably expressing empty vector (-), wild-type RAG-2 (wt), or the indicated RAG-2 mutation were arrested with STI-571 for 48 hours (control cells were treated with the same volume of DMSO). Recombination of the Igκ locus was assayed by southern blot as described Figure 10.

(B) Quantification of recombination at the Igκ locus.

(C) IgH recombination in R2K3 cells. Top: R2K3 cells were treated as in (A) and recombination of the IgH locus was assayed by southern blot as in Figure 10. Bottom: PCR of RAG1 visualized by ethidium bromide as a control for DNA quality.

(D) Quantification of recombination at the IgH locus.

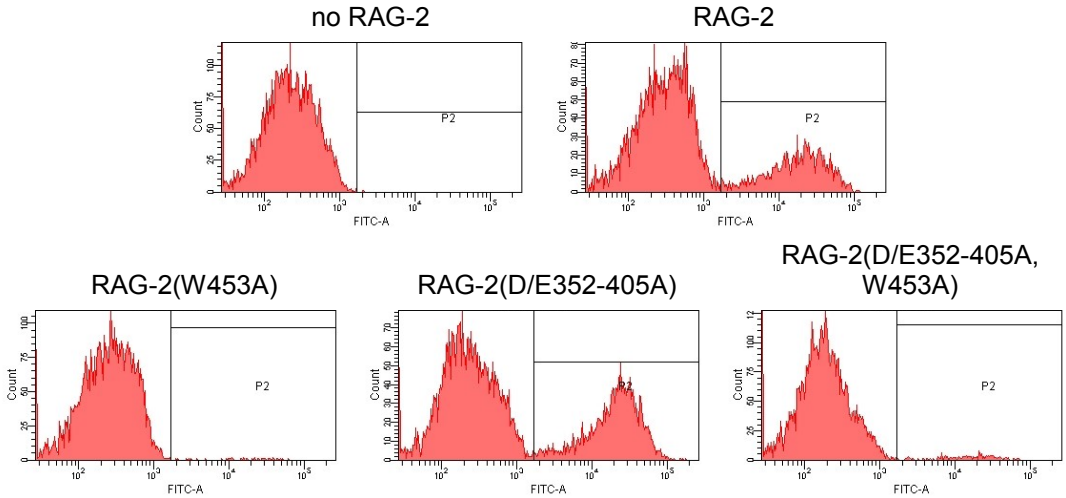
(E) Diagram of the IgH locus with the assayed germline transcripts noted. The μ0 transcript is initiated at a promoter 5' of DQ52, the most 3' D-segment. The PCR product detected for this transcript is noted as a line 3' of DQ52. Multiple Iμ transcripts originate from the Eμ enhancer 3' of the J-segments in the direction of the constant (Cμ) region. The approximate location of the PCR product is shown with a line.

(F) Assay for germline transcription at the IgH locus. RNA was isolated from independent R2K3 samples arrested with STI-571 for 48 hours. cDNA was generated with (+) and without (-) reverse-transcriptase (RT) and a PCR was performed for Iμ (top two panels), μ0 (middle two panels), and actin (ACTB) as a control (bottom two panels).

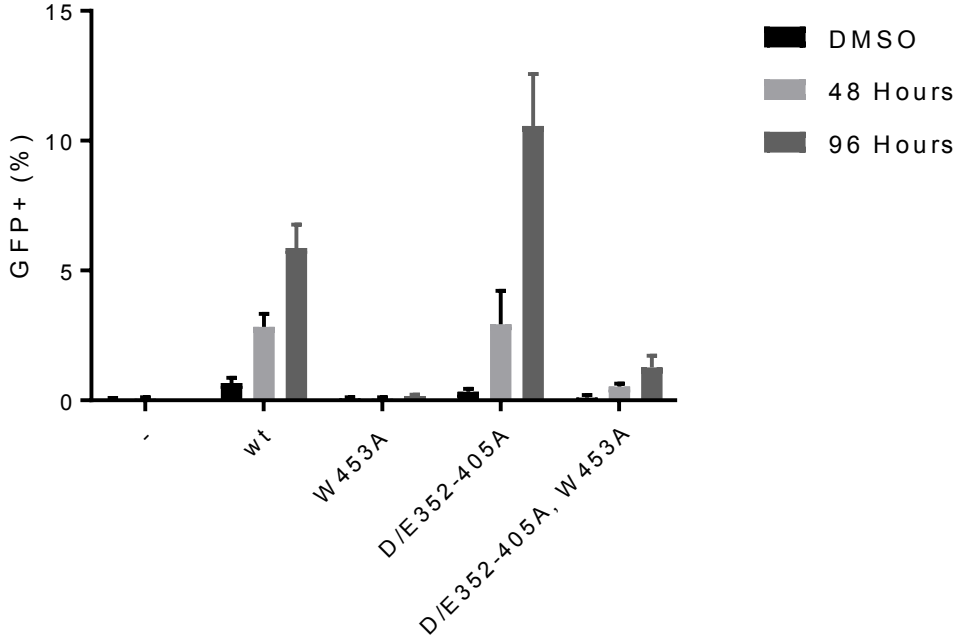
A


STI-571 G1 arrest
Inversional joining
GFP
GFP
GFP
RAG activity

B


no RAG-2
RAG-2
RAG-2(W453A)
RAG-2(D/E352-405A)
RAG-2(D/E352-405A, W453A)
Count
FITC-A
P2

C


GFP+ (%)
15
10
5
0
DMSO
48 Hours
96 Hours
-
wt
W453A
D/E352-405A
D/E352-405A, W453A

D

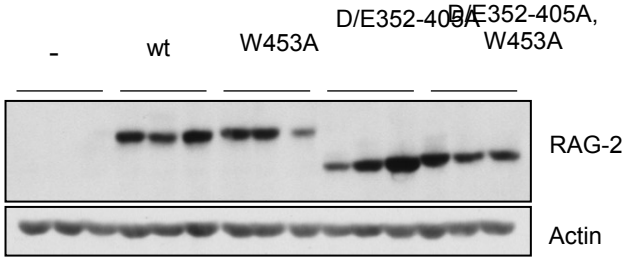

-
wt
W453A
D/E352-405A
D/E352-405A, W453A
RAG-2
Actin

Figure 12. The RAG-2 inhibitory domain suppresses recombination on an integrated substrate in R2K3 cells

(A) Diagram of the integrated substrate, PMX-INV (Gapud et al. 2011). An inverted GFP cassette is flanked by two RSSs (triangles) in the context of genomic sequence (represented as squares). After arrest with STI-571 induces arrest, RAG cleaves at the RSS heptamer and the reporter is inverted and joined so that GFP is constitutively expressed.

(B) Representative flow cytometry plots of GFP expression of the indicated RAG variant arrested with STI-571 for 96 hours.

(C) Quantification of flow cytometry for GFP expression for three independent infections (mean ± SD). Bars give the percentage of cells GFP positive in control treated (DMSO) cells and cells arrested with STI-571 for 48 or 96 hours.

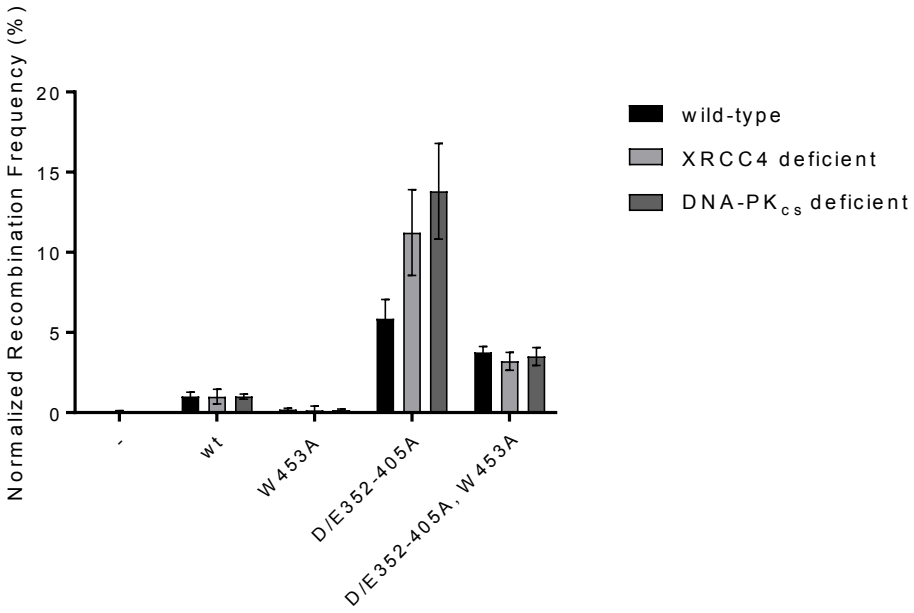
(D) Immunodetection of the myc-tagged RAG-2 variants transduced in R2K3 cells with actin below as a loading control.

The RAG-2 inhibitory domain acts prior to repair

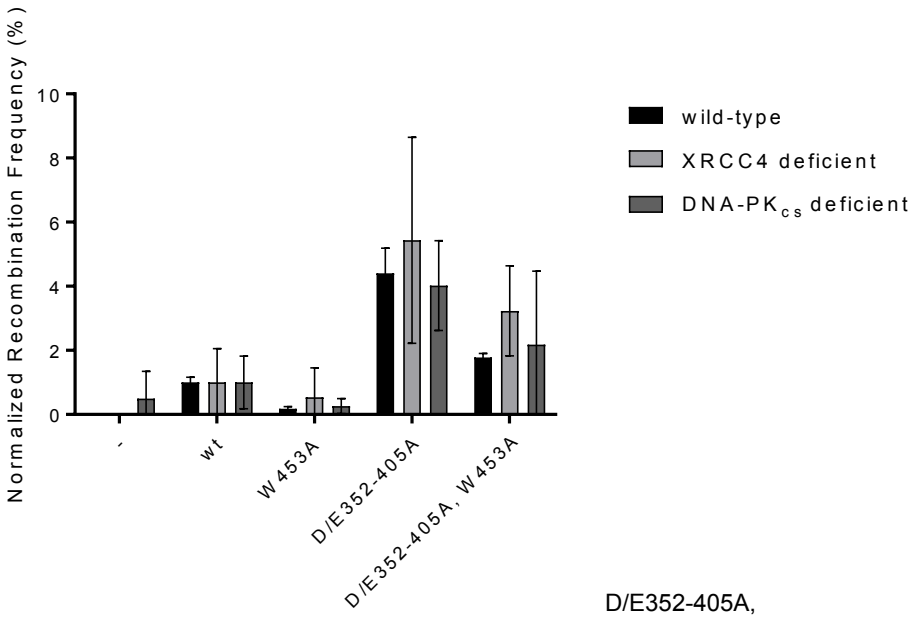
In the experiments described above we defined an inhibitory domain in RAG-2, the disruption of which stimulated V(D)J recombination within extrachromosomal and integrated substrates as well as at endogenous loci. RAG participates in all stages of V(D)J recombination from binding to synapsis and cleavage to repair. RAG-induced breaks are repaired in collaboration with components of the NHEJ machinery. We asked whether the suppressive effect of the inhibitory domain was exerted during repair. To do so, we employed cell lines deficient in the NHEJ factors XRCC4 or DNA-PK_{CS}. Should the inhibitory domain function through an interaction with the NHEJ machinery, then in the absence of NHEJ we would expect inactivation of the inhibitory domain to fail to stimulate recombination. In extrachromosomal assays for signal joint (Figure 13A) and coding joint (Figure 13B) formation the D/E352-405A mutation reversed the effect of the W453A mutation to a similar extent in NHEJ-proficient and NHEJ-deficient cell lines. Moreover, RAG-2(D/E352-405A) exhibited a robust gain-of-function, compared to wild-type in all three cell lines (Figure 13A, B). The effects of these mutations were not attributable to differences in protein expression (Figure 13C). Prior work suggested that mutation of acidic residues within this region increases the usage of alternative NHEJ (alt-NHEJ) (Coussens et al. 2013), which is characterized by excessive deletions and microhomology (Deriano & Roth 2013). The use of microhomology is not, however, restricted to alt-NHEJ (Pannunzio, Li, Watanabe, & Lieber, 2014). In wild-type cells, all but one of the recovered signal junctions produced by RAG-2 variants were precise (Figure 14), suggesting that the robust increase in recombination frequency associated with the D/E352-405 mutation does not result from the use of alt-NHEJ. In wild-type

cells, similar use of microhomology and deletion at coding joints were observed for all RAG-2 variants assayed (Figure 16E, F). While we could obtain few junctions in assays of wild-type RAG-2 in repair-deficient cell lines, the signal junctions obtained from cells expressing the RAG-2 inhibitory domain mutants were characterized by an increase in excessive deletion that was not accompanied by an increased use of microhomology (Figure 16A-D). Moreover, the relative increase in recombination activity conferred by disruption of the RAG-2 inhibitory domain is similar regardless of deficiencies in NHEJ. These data are therefore consistent with an increase in the activity of RAG-2(D/E352-405A) and RAG-2(D/E352-405A, W453A) at a step prior to repair.

A


Normalized Recombination Frequency (%)
20
15
10
5
0
-
wt
W453A
D/E352-405A
D/E352-405A, W453A
wild-type
XRCC4 deficient
DNA-PKcs deficient

B


Normalized Recombination Frequency (%)
10
8
6
4
2
0
-
wt
W453A
D/E352-405A
D/E352-405A, W453A
wild-type
XRCC4 deficient
DNA-PKcs deficient

C

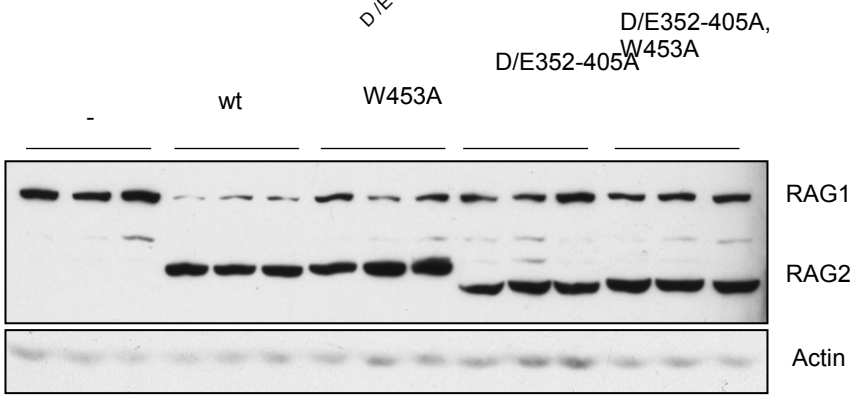

-
wt
W453A
D/E352-405A
D/E352-405A, W453A
RAG1
RAG2
Actin

Figure 13. The inhibitory domain suppresses recombination in the absence of classical non-homologous end joining

(A) Signal joining of RAG-2 inhibitory domain mutants in NHEJ-deficient cells. Extrachromosomal signal joining was assayed in wild-type Chinese hamster ovary (CHO) cells as well as derivatives deficient for XRCC4 and DNA-PK_{CS}. Signal joint recombination of each RAG-2 mutant was determined in three independent transfections in each of the three cell lines. Recombination frequency for each cell line was normalized to wild-type RAG-2 (wt). Plotted as normalized mean ± SD.

(B) Coding joining of RAG-2 inhibitory domain mutants in NHEJ-deficient cells. Assayed as in (A) with the pJH290, coding joint, substrate.

(C) Immunodetection of RAG-1 and RAG-2 in CHO cells. Top: anti-myc detection of RAG-1 and RAG-2. Bottom: actin detected as a loading control.


Plasmid TGTTTTTGTTCCAGTCTGTAGCACTGTG CAGGTCTC CTGAACCTG CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
Wild-type cells
wt (4/4) TGTTTTTGTTCCAGTCTGTAGCACTGTG CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N (5/5) TGTTTTTGTTCCAGTCTGTAGCACTGTG CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
DM (5/6) TGTTTTTGTTCCAGTCTGTAGCACTGTG CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
DM TGTTTTTGTTCCAGTCTGTAGCACTGTG --CAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
DNA-PKcs-deficient cells
352N (5/10) TGTTTTTGTTCCAGTCTGTAGCACTGTG CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCAC---- ---AGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCAC---- ---AGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCA----- ----GTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCACTGT- ----------CTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTG---------- A ---------AGTACTCCACTGTCTGGCTGTACAAAAACC
DM (2/5) TGTTTTTGTTCCAGTCTGTAGCACTGTG CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
DM TGTTTTTGTTCCAGTCTGTAGCAC---- ---AGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
DM TGTTTTTGTTCCAGTCTGTAGCACTGT- ----------ACAAAAACC
DM -125bp (entire RSS deleted)
XRCC4-deficient cells
352N TGTTTTTGTTCCAGTCTGTAGCACTGTG ----GTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCACTGTG -------GTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCAC---- ---AGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCNGTCTGTAG------- -----------TACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGC------ -------------CTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCT---------- ---------------CCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCACT--- ---------------CCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCAC---- T ----GTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCACTGTG CTT ---AGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
352N TGTTTTTGTTCCAGTCTGTAGCACTG-- CTGAACCTG CACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
DM TGTTTTTGTTCC---------- --CAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
DM TGTTTTTGTTCC---------- --CAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACC
DM TGTTTTTGTTCCAGTCTGTAGCACTGT- ---------AGTACTCCACTGTCTGGCTGTACAAAAACC
DM TGTTTTTGTTCCAGTCTGTAC------- ----------AAAAACC
DM -12bp ----------CTGTCTGGCTGTACAAAAACC
DM -12bp ----------GTCTGGCTGTACAAAAACC
DM TG---------- (24bp untemplated) -1
DM -300+bp (entire RSS deleted)
DM -300+bp (entire RSS deleted)

Figure 14. Signal joints from wild-type and NHEJ-deficient cells

Signal joints obtained from extrachromosomal assays with wild-type RAG-2 (wt), RAG-2(D/E352-405A) (352N), and the double mutant (DM), RAG-2(D/E352-405A, W453A) in wild-type, DNA-PK_{CS}-deficient, and XRCC4-deficient cells. The sequence of the plasmid is shown at the top with the heptamers indicated in blue and a portion of the intervening sequence shown. Signal joining is typically precise, with the flush heptamer of the 12-RSS joining to the flush heptamer of the 23-RSS. Due to the inability to rule out clonal expansion during bacterial transformation, only unique sequences from each individual transformation are counted. Each section lists the sequences of junctions from the given cell line, with the RAG variant that produced each junction on the left. Numbers in parentheses note the number of flush signal joints obtained out of the total number of unique joints. Bold typeface indicates microhomology and dashes represent deletions. The 24bp untemplated addition does not match any sequence in the plasmid or genome. The green type shows the templated insertion from the plasmid sequence between the two RSSs.

```
Plasmid     GCTGCAGGTCGAC                GGATCCCCGGGGATC
Wild-type cells
wt          GCTGCAGGTCGAC                GGATCCCCGGGGATC
wt          GCTGCAGGTCGAC GT             -----CCCGGGGATC
wt          GCTGCAGGTCGAC G              ---TCCCCGGGGATC
wt          GCTGCAGGTCGAC                -GATCCCCGGGGATC
wt          GCTGCAGGTCGA-                ---TCCCCGGGGATC
wt          GCTGCAGGTCGA-                ---TCCCCGGGGATC
wt          GCTGCAGGTCGAC                ----CCCCGGGGATC
wt          GCTGCAGGTCGAC                -----CCCGGGGATC
wt          GCTGCAGGTC---                -------CGGGGATC
W453A       GCTGCAGGTCGAC GT             -GATCCCCGGGGATC
W453A       GCTGCAGGTCGAC GT             -----CCCGGGGATC
W453A       GCTGCAGGTCGAC G              ---TCCCCGGGGATC
W453A       GCTGCAGGTCGAC                -GATCCCCGGGGATC
W453A       GCTGCAGGTCGA-                ---TCCCCGGGGATC
W453A       GCTGCAGGTCGA-                ---TCCCCGGGGATC
W453A       GCTGCAGGTCGAC                -----CCCGGGGATC
W453A       GCTGCAGGTC---                ------CCGGGGATC
352N        GCTGCAGGTCGAC GAC            GGATCCCCGGGGATC
352N        GCTGCAGGTCGAC GT             -----CCCGGGGATC
352N        GCTGCAGG----- AT             -----CCCGGGGATC
352N        GCTGCA------- T              -----CCCGGGGATC
352N        GCTGCAGGT----                GGATCCCCGGGGATC
352N        GCTGCAGG-----                --ATCCCCGGGGATC
352N        GCTGCAGGTCGA-                ---TCCCCGGGGATC
352N        GCTGCAGGTCGAC                ----CCCCGGGGATC
352N        GCTGCAGGTCG--                -------CGGGGATC
352N        GCTGCAGGTC---                -------CGGGGATC
DM          GCTGCAGGT---- GAT            -----CCCGGGGATC
DM          GCTGCAGGTCGAC GT             -----CCCGGGGATC
DM          GCTGCAGGTCGA- T              -----CCCGGGGATC
DM          GCTGCAGGTCGAC                -GATCCCCGGGGATC
DM          GCTGCAGGTCGAC                -GATCCCCGGGGATC
DM          GCTGCAGGCC---                -GATCCCCGGGGATC
DM          GCTGCAGGTCGAC                -----CCCGGGGATC
DM          GCTGCAGGTC---                ------CCGGGGATC
DM          GCTGCAGG-----                ------------ATC
DNA-PKcs-deficient cells
wt           CTGCAGGTCGAC GTCGACTGGATCC  GGATCCCCGGGGATC
352N        GCTGCAGGTCGAC GTCGATCC       GGATCCCCGGGGATC
352N        GCTGCAGGTCGAC GTC            -GATCCCCGACGGATC
352N         TGCAGGTCGAC GTC             ----CCCCGGGGATC
DM          GCTGCAGGTCGAC GT                      -22
XRCC4-deficient cells
wt           TGCAGGTCGAC GTC             -GATCCCCGGGGATC
352N        GCTGCAGGT----                -GATCCCCGGGGATC
352N        GCTGCAGGTC---                ------CCGGGGATC
352N        GCTGCAGGTCGAC GTCC           GGATCCCCGGGGATC
352N        GCTGCAGGTCGAC GTCC           GGATCCCCGGGGATC
352N        GCTGCAGGTCGAC GTC            GGATCCCCGGGGATC
352N        GCTGCAGGTCGAC GTC            -GATCCCCGGGGATC
352N        GCTGCAGGTCGA- TCC            GGATCCCCGGGGATC
352N        GCTGCAGGTCGAC GTCGATCC       GGATCCCCGGGGATC
352N        GCTGCAGGTCGAC GTCGATCC       GGATCCCCGGGGATC
352N        GCTGCA------- CTA                      -19
```

Figure 15. Coding joints from wild-type and NHEJ-deficient cells

Coding joints obtained from extrachromosomal assays with wild-type RAG-2 (wt), RAG-2(D/E352-405A) (352N), and the double mutant (DM), RAG-2(D/E352-405A, W453A) in wild-type, DNA-PK$_{CS}$-deficient, and XRCC4-deficient cells. The plasmid sequence flanking the RSSs is shown at the top. RAG cleavage results in coding hairpins that are opened by the Artemis endonuclease after DNA-PK$_{CS}$ phosphorylation. As in Figure 14, only unique sequences from each transformation reaction are counted. Sequences of complex translocations were obtained, but could not be succinctly described. These were included in the analyses of microhomology and deletions, but excluded from this figure. Bold typeface indicates microhomology and dashes represent deletions. Palindromic insertions, formed from Artemis opening the hairpin with a single strand nick, are underlined with potential microhomology in the palindromic sequences marked with a thick underline. Red nucleotides are mutated from the plasmid sequence.

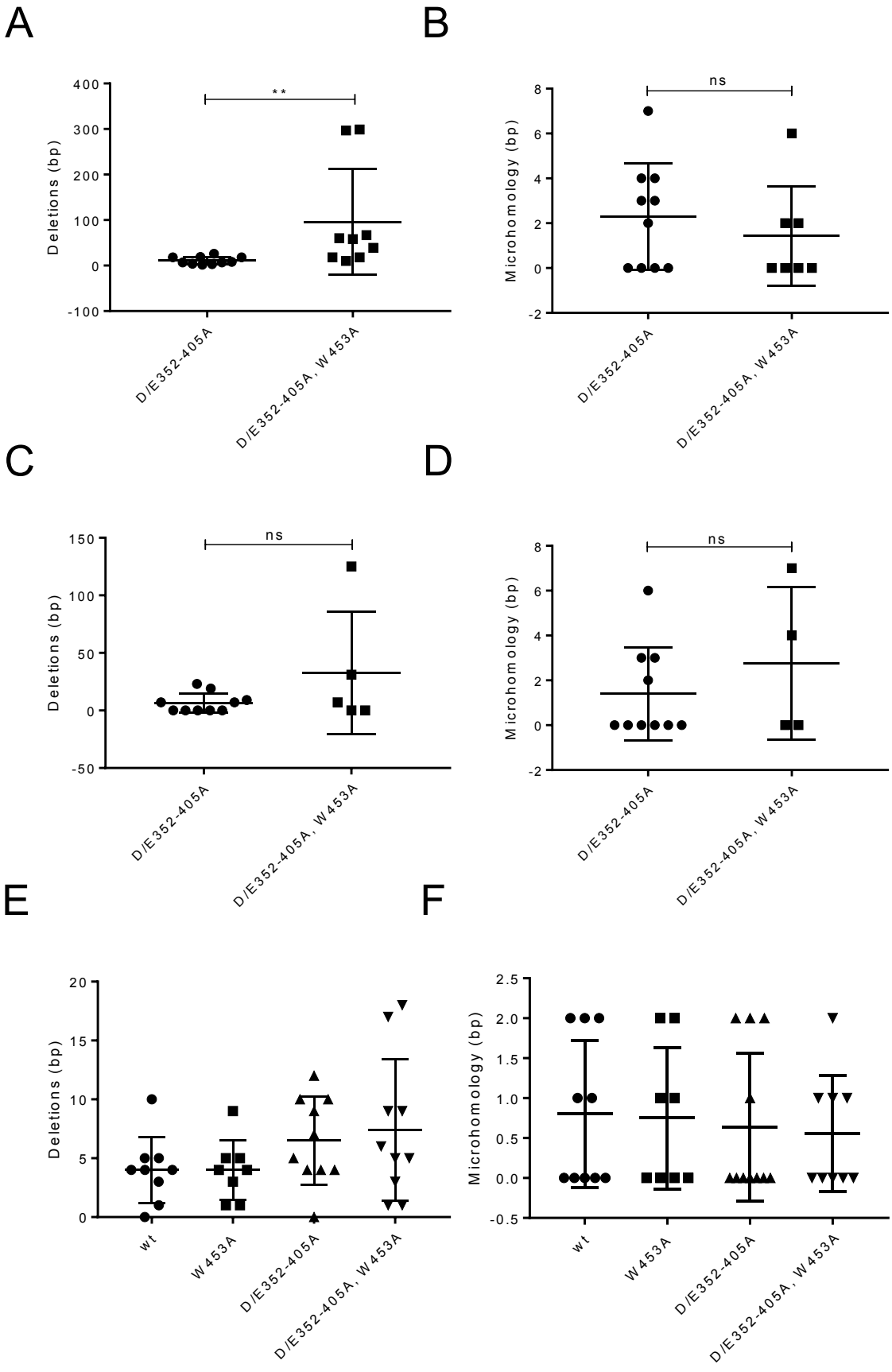

A
B
C
D
E
F
Deletions (bp)
Microhomology (bp)
**
ns
D/E352-405A
D/E352-405A, W453A
wt
W453A

Figure 16. Analysis of junctions sequenced junctions from NHEJ-deficient and control cells

(A) Distribution of the length of deletions in signal junctions from XRCC4-deficient cells. Presented mean ± SD. Kruskal-Wallis test was performed, ** indicates $p < 0.005$.

(B) Lengths of microhomology in signal junctions from XRCC4-deficient cells. Analyzed as in (A), ns indicates $p > 0.05$.

(C) Distribution of the length of deletions in signal joints from DNA-PK_{CS}-deficient cells analyzed as in (A), ns indicates $p > 0.05$.

(D) Lengths of microhomology of deletions in signal joints from DNA-PK_{CS}-deficient cells analyzed as in (A), ns indicates $p > 0.05$.

(E) Distribution of the length of deletions and (F) microhomology in coding junctions from wild-type cells. For both, a Kruskal-Wallis test showed the means were not significantly different $p > 0.05$.

The RAG-2 inhibitory domain gates access to chromatin (in collaboration with Gita Kumari and Ranjan Sen)

The distribution of RAG-2 over chromatin is positively correlated with the density of H3K4me3, both within and outside of antigen receptor loci (Ji et al. 2010). While recognition of H3K4me3 by RAG-2 is essential for efficient V(D)J recombination *in vivo* ((Liu et al. 2007) and Figures 3, 10, and 11), this requirement can be bypassed by inactivation of the inhibitory domain (Figures 3, 10, and 11). These observations, together with the robust stimulatory effect that disruption of the inactivating domain exerts on recombination, suggested that the RAG-2 inhibitory domain might function in modulating access of RAG to chromatin. To test this, we used chromatin immunoprecipitation (ChIP) to probe the distribution of RAG-2 and RAG-1 over immunoglobulin loci in 63-12 or R2K3 B-progenitor cells expressing wild-type RAG-2 or RAG-2 mutants.

In 63-12 cells, wild-type RAG-2 is detected at IgH, where it is localized primarily to DQ52 and the J_H cluster (Figure 17A), as well as over actively transcribed non-Ig loci (Figure 17A and Figure 18C, wt), consistent with previous observations (Ji et al. 2010). The RAG-2 W453A mutation, which abolishes binding to H3K4me3 (Liu et al. 2007), eliminates association of RAG-2(W453A) with the IgH locus and other active loci (Figure 17A and Figure 18C, W453A). Strikingly, disruption of the inhibitory domain in the context of the W453A mutation allows RAG-2(D/E352-405A, W453A) to access the IgH locus in the absence of H3K4me3 binding (Figure 17A, D/E352-405A, W453A), without regaining the ability to bind to adventitious sites (Figure 18C, D/E352-405A, W453A). This is consistent with the interpretation that off-target binding of RAG-2 to

chromatin is mediated principally if not wholly by H3K4me3. In 63-12 cells expressing wild-type RAG-2, RAG-1 is associated with IgH at DQ52 and the J_H cluster but unlike RAG-2, RAG-1 is not found at the ɣ-actin locus (Figure 17B, wt). Mutation of the RAG-2 PHD finger greatly reduces the binding of RAG-1 to the IgH locus, suggesting that this association is established or maintained by the interaction of H3K4me3 with RAG-2 (Figure 17B, W453A). Introduction of a second-site mutation in the RAG-2 inhibitory domain reverses the effect of the W453A mutation, allowing RAG-1 to bind the IgH locus in the absence of an interaction between RAG-2 and H3K4me3 (Figure 17B, 388-405A_{18}, W453A).

Similar patterns of RAG-2 and RAG-1 association with the IgH locus and of RAG-2 with active non-Ig loci were observed in R2K3 cells expressing wild-type RAG-2 (Figure 17C, D and Figure 18B, wt) and these patterns were unaffected by mutation of the inhibitory domain alone (Figure 17C, D and Figure 18B, D/E352-405A). This binding pattern correlated with H3K4me3 deposits at the IgH locus and did not differ between mutants (Figure 18A). As was observed for 63-12 cells, the RAG-2 W453A mutation greatly reduced binding of RAG-2(W453A) and RAG-1 to all regions of chromatin examined (Figure 17C, D and Figure 18B, W453A), while a second mutation in the RAG-2 inhibitory domain restored association of RAG-2(D/E352-405A, W453A) and RAG-1 with the IgH locus (Figure 17C, D and Figure 18B, D/E352-405A, W453A). RAG-2 and RAG-1 were associated with the J_κ cluster of the Igκ locus in R2K3 cells expressing wild-type RAG-2 (Figure 17E, F). While the W453A mutation reduced binding of RAG-2 to the Igκ locus (Figure 17E, W453A), the reduction was not as great as observed at the IgH locus; moreover, disruption of the inhibitory domain did not

reverse the effect of the W453A mutation on RAG-2 binding at the Igκ locus (Figure 17E, D/E352-405A, W453A). Association of RAG-1 with the Jκ cluster was unimpaired by the RAG-2 W453A mutation (Figure 17F, W453A), suggesting that the binding of RAG-1 to the Igκ locus is independent of H3K4me3 engagement by RAG-2. Taken together these observations indicate that the binding of RAG-1 and RAG-2 to the IgH locus is dependent, directly or indirectly, on engagement of H3K4me3 by the RAG-2 PHD finger, and that this requirement is imposed by the RAG-2 inhibitory domain. Disruption of the inhibitory domain allows RAG-2 and RAG-1 to access the IgH locus in a manner independent of H3K4me3 binding.

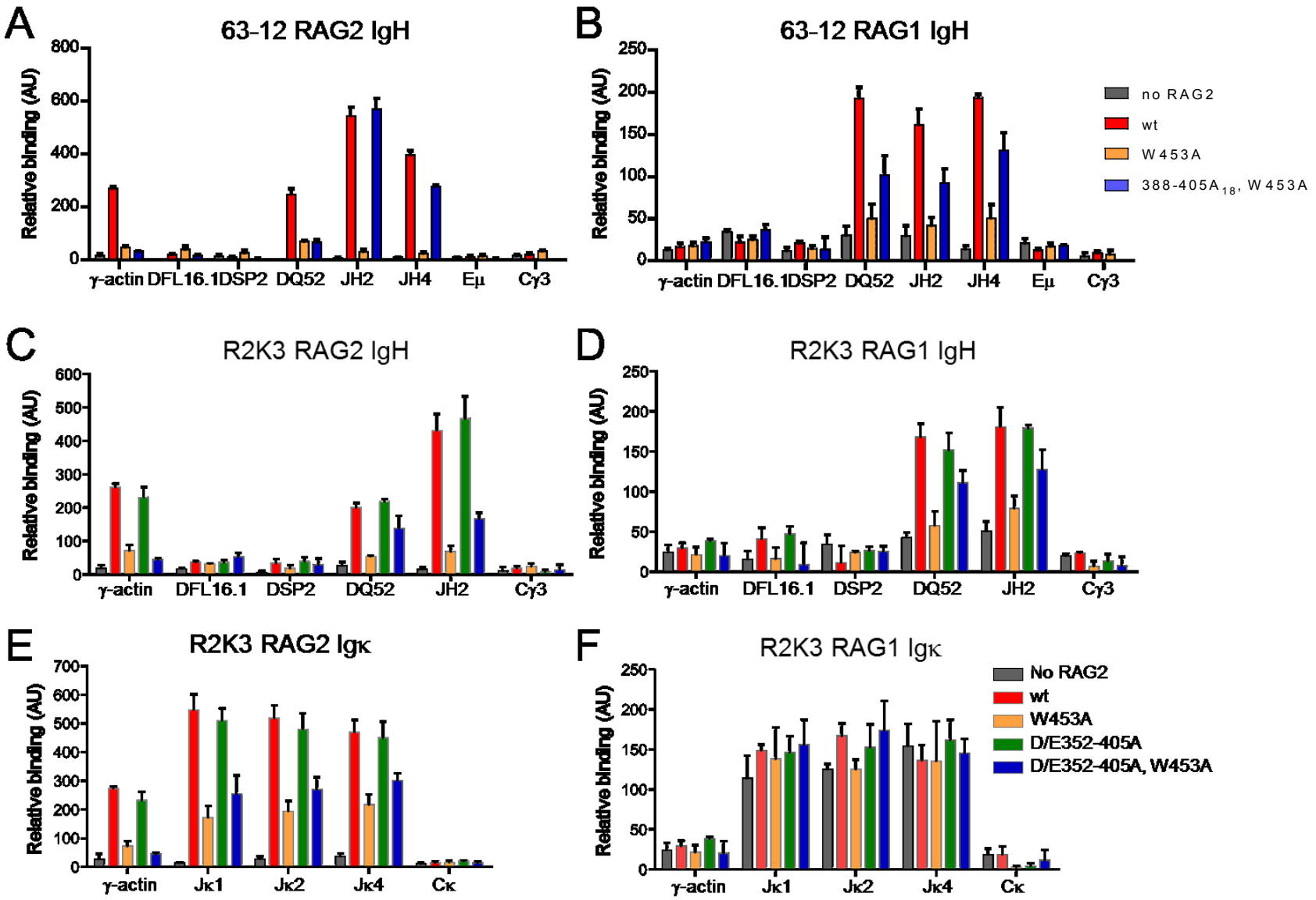

A
63-12 RAG2 IgH
Relative binding (AU)
800
600
400
200
0
γ-actin
DFL16.1
DSP2
DQ52
JH2
JH4
Eμ
Cγ3
B
63-12 RAG1 IgH
Relative binding (AU)
250
200
150
100
50
0
γ-actin
DFL16.1
DSP2
DQ52
JH2
JH4
Eμ
Cγ3
no RAG2
wt
W453A
388-405A18, W453A
C
R2K3 RAG2 IgH
Relative binding (AU)
600
500
400
300
200
100
0
γ-actin
DFL16.1
DSP2
DQ52
JH2
Cγ3
D
R2K3 RAG1 IgH
Relative binding (AU)
250
200
150
100
50
0
γ-actin
DFL16.1
DSP2
DQ52
JH2
Cγ3
E
R2K3 RAG2 Igκ
Relative binding (AU)
700
600
500
400
300
200
100
0
γ-actin
Jκ1
Jκ2
Jκ4
Cκ
F
R2K3 RAG1 Igκ
Relative binding (AU)
250
200
150
100
50
0
γ-actin
Jκ1
Jκ2
Jκ4
Cκ
No RAG2
wt
W453A
D/E352-405A
D/E352-405A, W453A

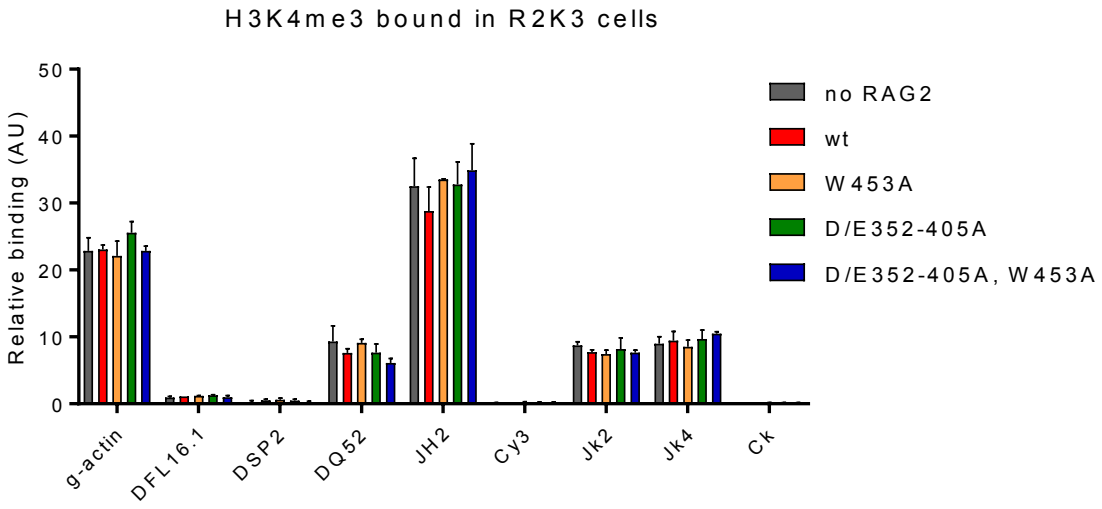
Figure 17. The RAG-2 inhibitory domain regulates access to the IgH locus

(A) Chromatin immunoprecipitation (ChIP) of RAG-2 and (B) RAG-1 at the IgH locus in 63-12 cells expressing wild-type (wt) RAG-2 or the indicated variant. Enrichment of RAG was assayed by qPCR at the indicated regions as described (Ji et al. 2010). The constant region, Cɣ3 served as a negative control, while ɣ-actin served as a positive control for RAG-2, which binds promiscuously to H3K4me3 sites (Liu et al. 2007). Bars represent mean ± SD, n=2.

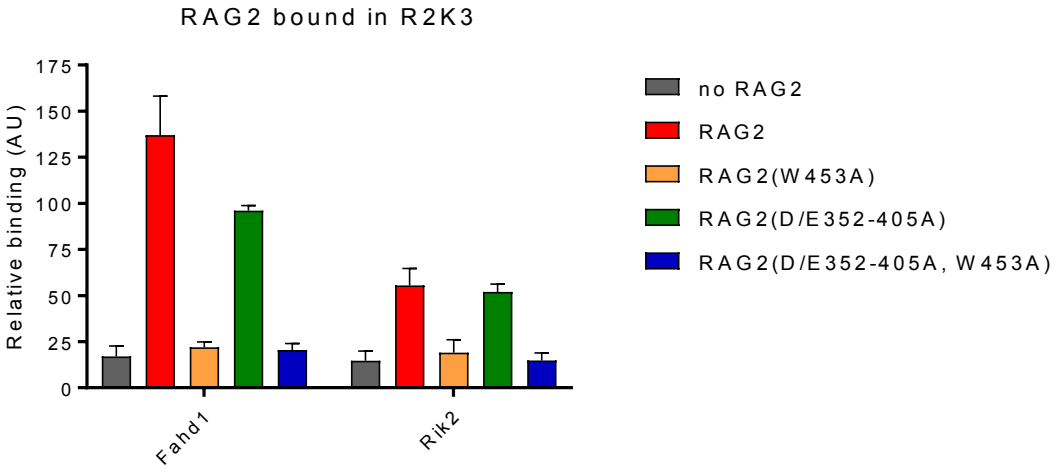
(C) ChIP of RAG-2 and (D) RAG-1 at the IgH locus in R2K3 cells expressing wild-type (wt) RAG-2 or the indicated variant as in (A).

(E) ChIP of RAG-2 (E) and RAG-1 (F) at the Igκ locus in R2K3 cells expressing wild-type (wt) RAG-2 or the indicated variant as in (A).

A


H3K4me3 bound in R2K3 cells
Relative binding (AU)
50
40
30
20
10
0
g-actin
DFL16.1
DSP2
DQ52
JH2
Cy3
Jk2
Jk4
Ck
no RAG2
wt
W453A
D/E352-405A
D/E352-405A, W453A

B


RAG2 bound in R2K3
Relative binding (AU)
175
150
125
100
75
50
25
0
Fahd1
Rik2
no RAG2
RAG2
RAG2(W453A)
RAG2(D/E352-405A)
RAG2(D/E352-405A, W453A)

C

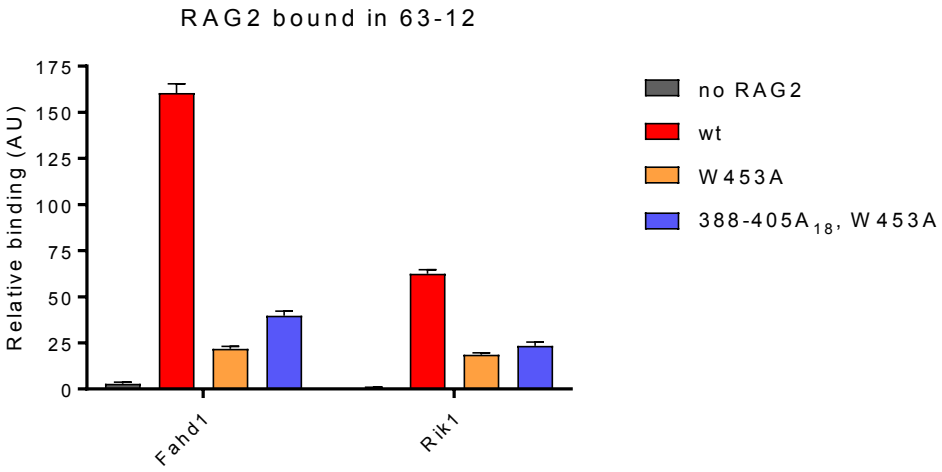

RAG2 bound in 63-12
Relative binding (AU)
175
150
125
100
75
50
25
0
Fahd1
Rik1
no RAG2
wt
W453A
388-405A18, W453A

Figure 18. RAG-2 with an intact PHD binds to sites of H3K4me3

(A) ChIP detection of H3K4me3 at the IgH locus in R2K3 cells. Enrichment of H3K4me3 was assayed by qPCR at the indicated regions as in Figure 17. The actively transcribed gene, ɣ-actin served as a positive control, Bars represent mean ± SD, n=2. (B) ChIP detection of RAG-2 binding to adventitious H3K4me3-dense sites in R2K3 and (C) 63-12 cells. Assayed as in Figure 17.

Discussion

Our observations support a model in which the responsiveness of RAG to epigenetic stimulation is conferred by a PHD-mediated inhibitory domain (PID) whose action is relieved upon binding of H3K4me3. The ability of exogenous H3K4me3 to stimulate the coupled cleavage activity of wild-type RAG was dependent, as expected, on binding of H3K4me3 to the PHD finger of RAG-2. H3K4me3 binding exerts at least two effects that contribute to enhanced RSS cleavage activity *in vitro*: increased affinity for substrate and faster catalysis. The accessibility of antigen receptor loci to RAG is associated with epigenetic modifications characteristic of active chromatin, such as H3K4me3, whose recognition by RAG-2 promotes V(D)J recombination. The ability of H3K4me3 to stimulate purified RAG cleavage *in vitro* suggests that H3K4me3 relieves inhibition exerted by some feature of the RAG complex. While core RAG-2 supports V(D)J recombination *in vitro* and *in vivo* in the absence of a PHD, RAG-2 bearing the point mutation W453A supports recombination *in vitro* and exhibits defects in recombination *in vivo*. This suggests that H3K4me3 is not minimally required for recombination, it is a requirement that is imposed by noncore RAG-2. Our results are consistent with a model in which this PID maintains RAG in a state of low affinity for RSSs until allosteric activation through the deposition of H3K4me3.

We initially identified an acidic inhibitory domain (AcID) whose mutation could bypass the impairment imposed by inactivation of the PHD. Disruption of this inhibitory region uncouples V(D)J recombination from the requirement for H3K4me3 binding by RAG-2. Interestingly, mutation of the AcID mimics the binding of H3K4me3 by increasing the affinity of RAG for substrate and enhancing its catalytic rate. While RAG-2(388/405A_{18}) supports a basal affinity for substrate that is also similar to the maximal

induced affinity of wild-type RAG, H3K4me3 induces no further increase in affinity. Affinity for substrate was nominally equal between wild type RAG-2 with H3K4me3 and the mutant RAG-2(388/405A_{18}) both with H3K4me0 and H3K4me3. We imagine several possible reasons that the affinity for substrate of the 388/405A_{18} cannot be stimulated by H3K4me3. It is possible that the observed K_D represents the maximum binding capabilities of the RAG complex. The observed K_D could also represent the maximum affinity that can be detected by the assay. Third, it is possible that the inhibitory domain relieved by H3K4me3 and the AcID increase affinity for substrate through the same mechanism, thus they have the same effect on K_D. The basal catalytic rate supported by RAG-2(388/405A18) is similar to the maximally induced rate observed for wild-type RAG, but is further increased in response to H3K4me3. Thus, the effects of H3K4me3 on substrate affinity and catalysis are separable. While the AcID mutation rescued activity of the RAG-2(W453A) mutant *in vivo*, and mimicked binding of H3K4me3 *in vitro*, subsequent study showed that it represents a separate domain that functions, at least in part, independent of the PHD.

The boundaries of the AcID lie within an acidic region of RAG-2, comprising residues 352-405. Neutralization of charge in this interval is associated with aberrant repair of RAG-mediated DSBs, decreased stability of RAG-signal end complexes, and genomic instability (Coussens et al. 2013). While these effects appear to reflect events occurring after RSS recognition and DNA cleavage, they may be explained in part by our results. An increase in genomic instability, for example, would be consistent with the H3K4me3-independent binding we observe upon mutation of the AcID. A unifying hypothesis would suggest that the destabilization of signal end complexes and relaxation

of repair pathway choice are consequences of the structural alterations that uncouple RAG activity from H3K4me3 binding upon mutation of the AcID. For example, if formation of a stable signal end complex were to require disengagement of RAG from H3K4me3, then mutations that mimic the effect of H3K4me3 engagement, such as those identified herein, could compromise signal end complex stability and appropriate repair of DNA ends.

Available data do not provide structural insights into either of these modes of inhibition. Recently, the crystal structure of RAG-2 core, from residue 1 to 387, was solved. However, there was no density detected for the region from 352 to 387. This is consistent with our boundary for the AcID, however it suggests that the inhibition exhibited in the truncation mutants is not caused by a protein-protein interaction that would have stabilized the AcID. Similarly, because the crystal structure of the RAG-2 PHD and core were solved separately, there is no structure that provides information about how the signal may be propagated. To address this, pulse alkylation mass spectrometry was used to determine which cysteines were exposed on the RAG complex in the presence of H3K4me3 or H3K4me0 (Bettridge et al. 2017). Robust allosteric changes in RAG-1 were observed in complexes in which H3K4me3 is bound to RAG-2. Specifically, there were reproducible changes in the DDBD of RAG-1 that may explain the increased affinity for substrate observed in wild-type RAG-2 in the presence of H3K4me3. Our results are consistent with a model in which a PID maintains RAG in a state of low affinity for RSSs until nearby transcriptional activation promotes allosteric activation through the deposition of H3K4me3. Due to the similarity of the observed K_Ds for protein complexes relieved of PID- or AcID-mediated inhibition, it is possible that

both inhibitory domains induce a conformational change in the DDBD. Further structural and pulse-alkylation experiments could provide insight into how these inhibitory domains function and how the inhibition imposed by the AcID is relieved.

Deletions longer than 100 bp were only formed in the presence of the RAG-2 double mutant, not in the presence of the single AcID mutant. Therefore, it is possible that the recognition of H3K4me3 contributes to fidelity of repair through a yet unknown mechanism. There is evidence that noncore RAG-2 contributes to repair of RAG-mediated breaks. Mutation of the acidic residues to alanines conferred a large gain of function to wild-type RAG-2. While these mutations neutralized the charge of these residues, they also decreased the steric constraints by decreasing the size of the side chains. Future work might test whether the neutralization of the aspartic acid and glutamic acid side chains in the AcID with asparagine and glutamine, respectively, results in the same relief of inhibition as neutralization with alanine. Recombination frequency was observed to decrease upon a much smaller replacement mutant of this type (Silver et al. 1993).

Our results are consistent with the possibility that the AcID exerts its suppressive effect through interactions with one or more regions of RAG distinct from the PHD finger. Due to the relief of inhibition of the RAG-2(W453A) construct through a second site AcID mutation, it is plausible that AcID mutation acts through a general increase of activity that is independent of the PHD. Mutation of the acidic inhibitory domain additionally confers the ability to bypass H3K4me3-dependent localization of RAG-2. The RAG-2 AcID appears to govern interactions with chromatin, as it allows correct localization of the RAG-2(W453A) mutant that is otherwise not capable of localization at

the IgH or Igκ loci. As RAG-2(W453A) behaves like wild-type RAG-2 *in vitro*, the inability to localize correctly *in vivo* likely contributes to the impairment on endogenous loci. The mechanisms that govern locus accessibility during V(D)J recombination have not been fully elucidated. Sterile germline transcription precedes recombination (Bolland et al. 2007) and is associated with establishment of a permissive chromatin state. Changes in the patterns of both H3K4me3 (Ji et al. 2010) and H3K27Ac (McMurry & Krangel 2000) are observed during development. The pattern of RAG-2 binding follows that of H3K4me3 (Ji et al. 2010), while RAG-1 appears to have one mode of binding governed by H3K4me3 and one that correlates with H3K27Ac (Maman et al. 2016). The H3K4me3 binding may be stronger in the presence of H3R2me2s (Ramón-Maiques et al. 2007). Further, DNA demethylation has been observed in a monoallelic distribution (Mostoslavsky et al. 1998) and specifically at the junctions of recombined alleles (Selimyan et al. 2013). Moreover, the initial reports of histone binding activity identified regions of the AcID important for histone binding, so this domain might function through binding to H3R2me2s, H3K27Ac, or a yet unidentified feature of RAG-accessible chromatin.

We now provide evidence that this acidic inhibitory domain (AcID) in RAG-2 gates access to chromatin in a manner independent from recognition of H3K4me3. When inhibition was relieved by the 388/405A$_{18}$ mutation, basal cleavage activity in the absence of H3K4me3 was similar to that observed for wild-type RAG in the presence of saturating H3K4me3 peptide. Similarly, the 388/405A$_{18}$ mutant exhibited the ability to bypass the requirement for H3K4me3 *in vivo* for localization of RAG-2 to IgH and Igκ and for localization of RAG-1 to IgH. The ability of RAG-1 to localize to all antigen

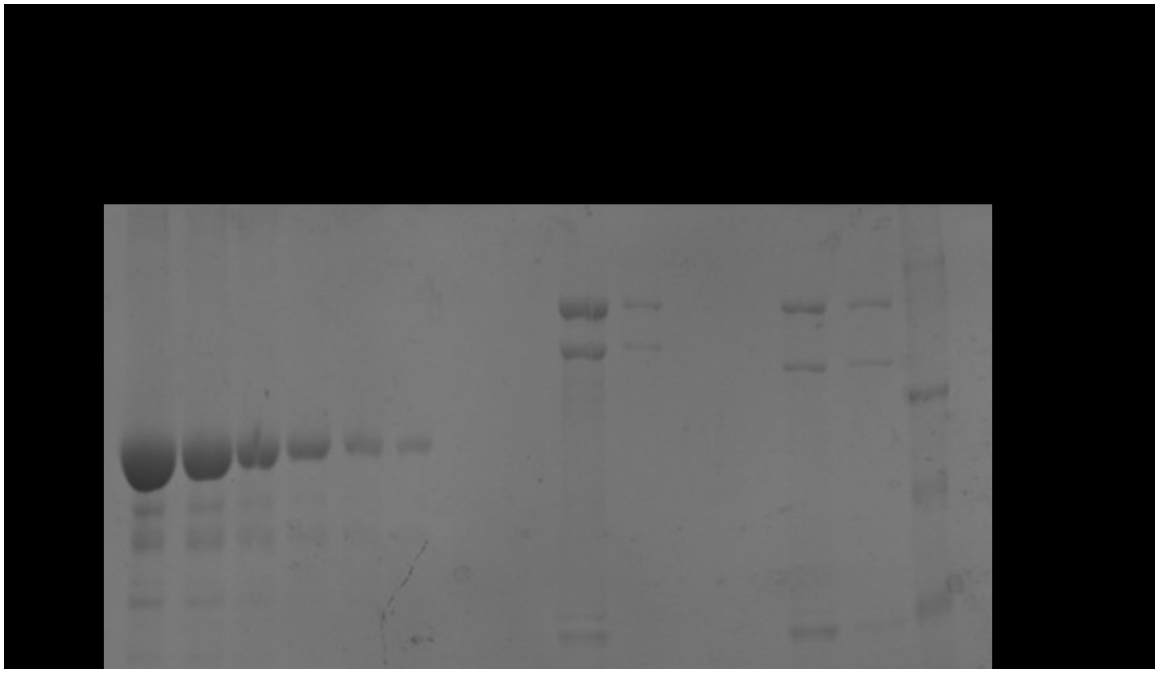
receptor loci aside from the IgH locus in the absence of RAG-2 may explain why Omenn syndrome patients can create T-cells, but have an almost complete block in B-cell development.

Mutation of AcID results in a robust gain of function *in vitro* and on extrachromosomal substrates. However, there is only a mild increase at endogenous loci and the integrated substrate data were equivocal. Similarly, the AcID mutation increased activity of RAG-2(W453A) to wild-type levels, or above, on extrachromosomal substrates, but recovered relatively less on recombination at the endogenous loci and on the integrated substrate. Previous work established that the extrachromosomal substrate is associated with H3K4me3 (Liu et al. 2007), so this difference suggests that a feature of endogenous loci is less permissive to this inhibitory domain mutant. Conversely, the extrachromosomal substrate may be more permissive than endogenous chromatin to recombination. This difference may reflect one or more of the still elusive additional mechanisms that govern locus accessibility to V(D)J recombination.

Supplement: Biochemical studies of the acidic inhibitory domain mutation in the context of RAG truncation

Purification of active RAG tetramers

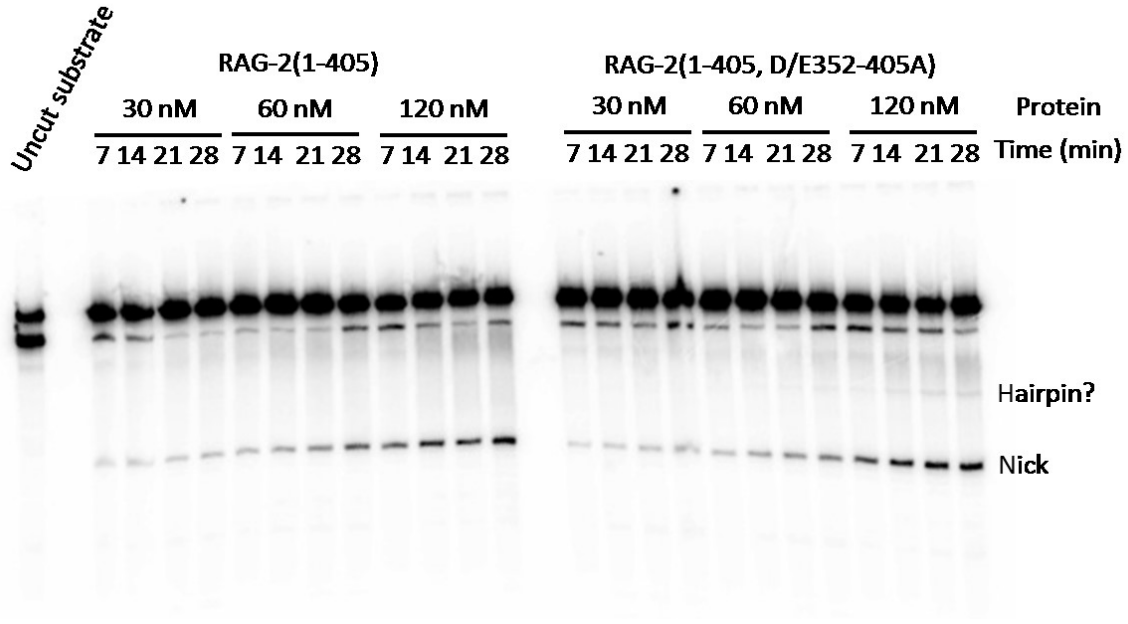
Our initial biochemical characterization of the RAG-2(388/405A_{18}) mutation was conducted in the presence of both the intact and inoperative PHD. Given our observations that the acidic inhibitory domain (AcID) functions independent of the PHD *in vivo* (Figure 9), we sought to extend these findings *in vitro*. To this end, we purified complexes of cRAG-1ct with truncated RAG-2 bearing an intact or neutralized AcID, RAG-2(1-405) and RAG-2(1-405, D/E352-305A) respectively. Both proteins eluted primarily in fraction 2 and the characteristic shift to lower apparent molecular weight seen in all neutralization mutants was evident (Figure S1A). We performed burst kinetics of pre-bound complexes at increasing nominal RAG concentrations over time (Figure S2). Interestingly, reactions containing the highest concentration of RAG bearing the AcID mutation was associated with hairpin product formation despite the omission of 23-RSS from these reactions (Figure S2A, 120 nM). The calculated active fractions of 11% for RAG-2(1-405) and 18% for RAG-2(1-405, D/E352-305A) were higher than previously observed (Figure 5). The curve slope is dictated largely by the highest points, so it is possible that using lower nominal RAG concentrations could change the observed values.



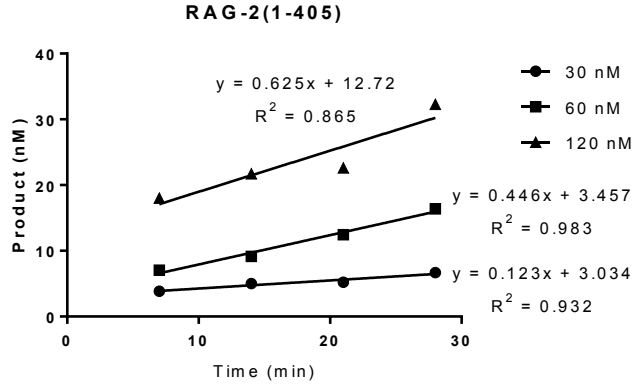
Supplemental Figure 1. Purification and normalization of activity of RAG truncations

Complexes of cRAG-1ctMH (cR1ct-MH) with RAG-2(1-405) or RAG-2(1-405, D/E352-405A) were purified by sonication and amylose affinity chromatography. Aliquots of each preparation were fractionated by SDS-PAGE alongside a dilution series of bovine serum albumin (BSA). Wash shows the wash of the amylose column; F1-F3 designate sequential amylose elution fractions. The positions of the RAG-1 and RAG-2 fusion proteins are indicated by black and gray arrowheads, respectively.

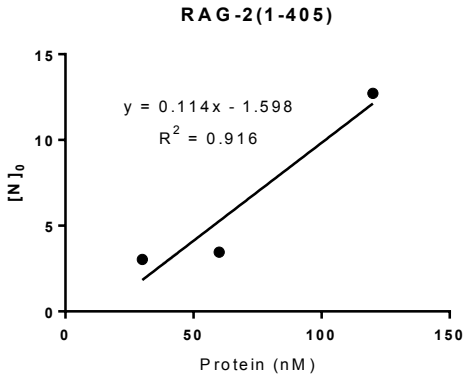
A


Uncut substrate
RAG-2(1-405)
30 nM
60 nM
120 nM
7 14 21 28 7 14 21 28 7 14 21 28
RAG-2(1-405, D/E352-405A)
30 nM
60 nM
120 nM
7 14 21 28 7 14 21 28 7 14 21 28
Protein
Time (min)
Hairpin?
Nick

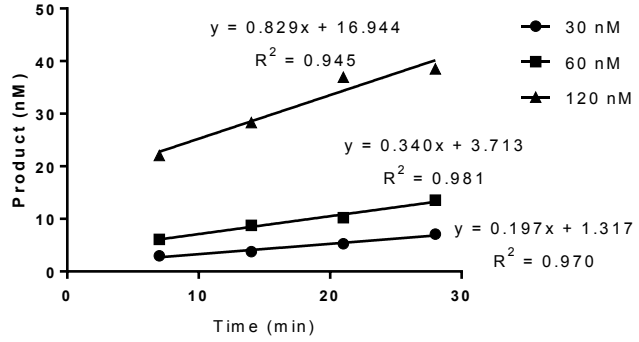
B


RAG-2(1-405)
y = 0.625x + 12.72
R2 = 0.865
y = 0.446x + 3.457
R2 = 0.983
y = 0.123x + 3.034
R2 = 0.932
30 nM
60 nM
120 nM
Product (nM)
Time (min)

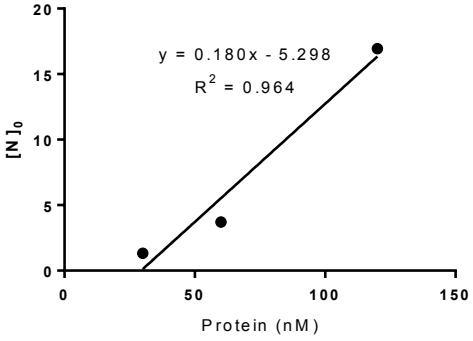
C


RAG-2(1-405)
y = 0.114x - 1.598
R2 = 0.916
[N]0
Protein (nM)

D


RAG-2(1-405, D/E352-405A)
y = 0.829x + 16.944
R2 = 0.945
y = 0.340x + 3.713
R2 = 0.981
y = 0.197x + 1.317
R2 = 0.970
30 nM
60 nM
120 nM
Product (nM)
Time (min)

E


RAG-2(1-405, D/E352-405A)
y = 0.180x - 5.298
R2 = 0.964
[N]0
Protein (nM)

Supplementary Figure 2. Burst kinetic determination of protein active fractions

(A) The indicated RAG-2 variant, copurified with cRAG-1ct, were assayed for nicking of a 12-RSS substrate at various nominal RAG concentrations. Reactions containing the indicated nominal RAG concentrations calculated on the basis of tetrameric stoichiometry were incubated for the indicated time. Uncut substrate shows the position of the substrate in the absence of RAG. The location of the Nicked product is indicated and Hairpin? Is used to point out a secondary product that roughly corresponds to the position of a hairpin product that should not have been formed in this assay. Product calculated by densitometry.

(B) Accumulation of nicked product as a function of time graphed from the image in (A). Graph of hairpin product (nM) accumulation over time at 30, 60, and 120 nM cRAG-1ct/RAG-2(1-405). Linear regression analysis was used to calculate nM product at the zero time point (t_0), with the given correlation coefficient (R^2).

(C) Estimation of active fraction of cRAG-1ct/RAG-2(1-405) complexes. The kinetic curves in (B) were extrapolated to zero time to give the $[N]_0$ associated with each nominal RAG concentration. $[N]_0$ was then plotted as a function of nominal RAG concentration; the slope of each resulting curve represents the fraction of active RAG in the preparation.

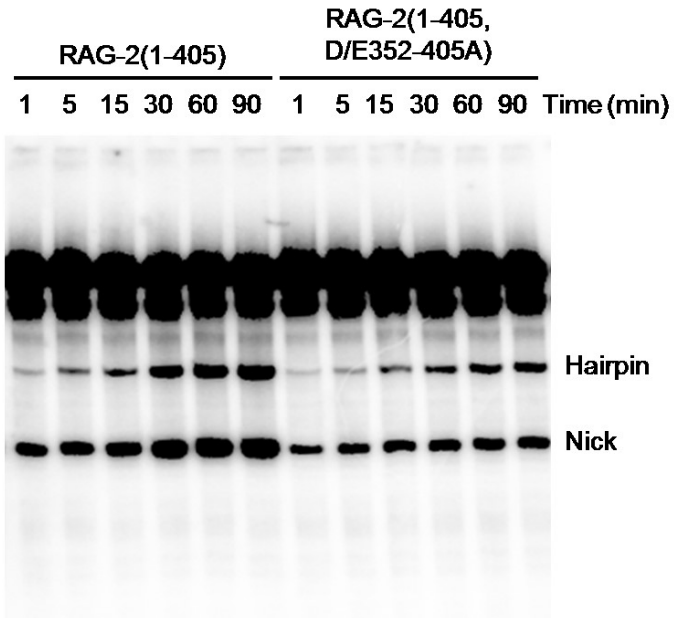
(D) Quantification of cRAG-1ct/RAG-2(1-405, D/E352-405A) as in (B).

(E) Graph of initial product formation at a given protein concentration for complexes of cRAG-1ct/RAG-2(1-405) as in (C).

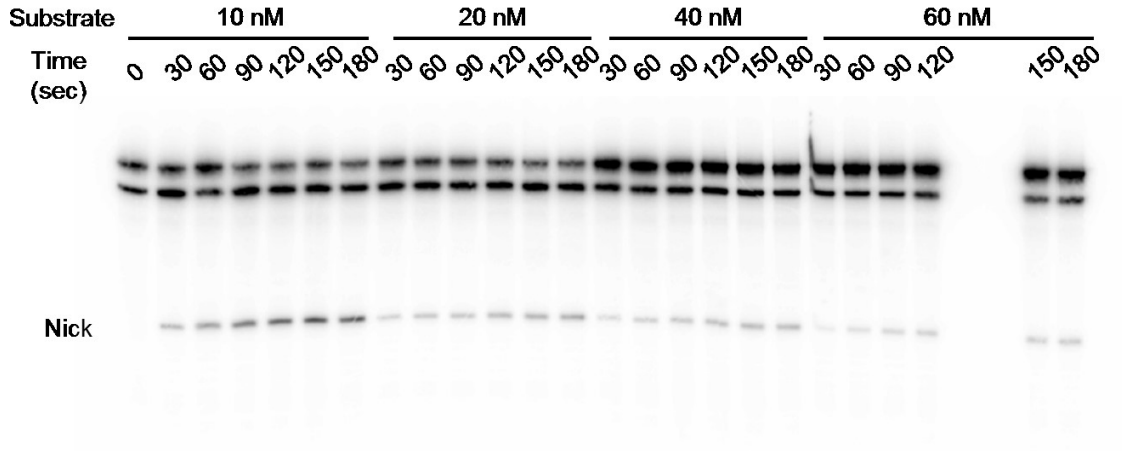
Biochemical analyses of acidic inhibitory domain mutation in the context of truncated RAG-2

Coupled cleavage assays showed more robust hairpin and nicked product formation in the RAG-2(1-405) complex than in the RAG-2(1-405, D/E352-405A) (Figure S3A). Similarly, nick kinetics of RAG-2(1-405) show more robust product formation that RAG-2(1-405, D/E352-405A) (Figure S3B, C). Additionally, as the active fraction of the AcID-mutated sample was higher, the assay was repeated without normalization for active fraction and the results remained the same. This leaves at least two plausible hypotheses: (1) The gain-of-function phenotype seen on the extrachromosomal assays with RAG-2(1-405, D/E352-405) is mediated by recognition of a factor that is not found in the biochemical reaction; (2) The biochemistry of this mode of inhibition is synthetically incompatible with the standard biochemical assays used to assay RAG activity. The first possibility is exciting in that it would narrow the possibilities of the elements that the AcID could be recognizing. It is however, unlikely as the original 388/405A_{18} mutation was characterized in the context of the full-length protein. The second possibility is more difficult to assess. Future biochemical experiments could use the 388/405A_{18} mutation in the context of RAG truncation and the D/E352-405A mutation in the context of full-length RAG-2 to determine whether the mutation or the truncation or the combination lead to decreased activity *in vitro*.

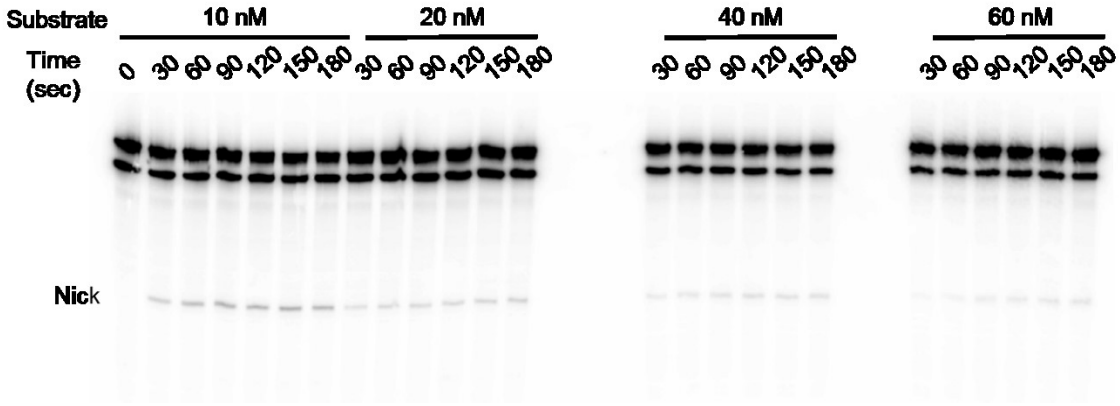
A


RAG-2(1-405)
RAG-2(1-405, D/E352-405A)
1 5 15 30 60 90 1 5 15 30 60 90 Time (min)
Hairpin
Nick

B


Substrate
10 nM
20 nM
40 nM
60 nM
Time (sec)
0 30 60 90 120 150 180 30 60 90 120 150 180 30 60 90 120 150 180 30 60 90 120 150 180
Nick

C


Substrate
10 nM
20 nM
40 nM
60 nM
Time (sec)
0 30 60 90 120 150 180 30 60 90 120 150 180 30 60 90 120 150 180 30 60 90 120 150 180
Nick

Supplementary Figure 3. Biochemical analysis of the acidic inhibitory domain mutation in the absence of noncore RAG-2

(A) Coupled cleavage. Complexes of cRAG-1ct and the indicated RAG-2 were incubated in the presence of radiolabeled 12-RSS and unlabeled 23-RSS substrate to assess coupled cleavage over time as indicated. Nick indicates the position of the nicked product, while Hairpin indicates the position of the hairpin product.

(B) Assay for RSS nicking. Reactions contained 1.5 nM cRAG-1ct/RAG-2(1-405) and 12-RSS substrate HL44/45 at 10, 20, 40, or 60 nM. Accumulation of nicked product (Nick) was assayed at the indicated times.

(C) Nick kinetics were assessed for cRAG-1ct/RAG-2(1-405, D/E352-405A) as in (B).

Materials and Methods

Cell culture

NIH3T3 and HEK 293T cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS) and 1X Penicillin-Streptomycin-Glutamine (PSG). R2K3 and 63-12 lines were propagated in RPMI1640 supplemented with 10% FBS, 50 μM 2-mercaptoethanol, 1X PSG, 0.7X MEM non-essential amino acids solution, 1 mM sodium pyruvate, 10 mM HEPES. R2K3 cells with integrated PMX-INV recombination substrate were kindly provided by Dr. Barry Sleckman (Weill Cornell Medicine) (Gapud et al. 2011). All cells were maintained at 37°C in 5% CO_2.

Antibodies

Commercial antibodies against the following proteins were used in this study: actin (clone AC-40, Sigma Aldrich, Cat#A3853); c-myc (clone 9E10, Santa Cruz Biotechnology, Cat#sc-40); RAG-1 (Abcam, Cat#ab172637); H3K4me3 (Active motif, cat#39915); sheep anti mouse secondary (GE Healthcare, Cat#NA931). The RAG-2 antibody was kindly provided by Dr. David Schatz (Yale University). Immunoblot incubations were performed in 5% nonfat dry milk in PBST (1X PBS, 0.1% Tween).

Expression constructs

RAG-1 and RAG-2, tagged at the N terminus with maltose-binding protein (MBP) tag and at the C-terminus with a c-myc epitope and a polyhistidine sequence, were expressed in the vectors pcDNA1 and pcDNA3.1 respectively. Methods and oligonucleotide sequences used to generate RAG-2 variants are given in Table 1. For GeneArt String (Thermo Fisher) constructs, endogenous PasI and PflMI restriction sites were used.

Truncation mutations were introduced by PCR using the forward primer indicated as RAG-2 truncation F and appropriate reverse primers followed by cleavage with BamHI and EcoRI.

Table 1. Cloning methods and oligonucleotide reagents for generation of RAG-2 variants

Mutation	Cloning Method	Primers 5' to 3'
RAG-2(P372A)	Site Directed Mutageneis	F: TGGAGTCACCAGCATCTTCTGTTGATGTCTGACT R: AGTCAGACATCAACAGAAGATGCTGGTGACTCCA
RAG-2(P377A)	Site Directed Mutageneis	F: TGAGTCTTCAAAGGCAGTGGAGTCACCAGGATC R: GATCCTGGTGACTCCACTGCCTTTGAAGACTCA
RAG-2(P372, 377A)	Site Directed Mutageneis	Sequential use of primers for P372A and P377A
RAG-2(334-351A_{18})	Divergent PCR	F: GCCGCTGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGAAGAGGATTTGAGTGAAG R: GCCGCTGCAGCGGCGGCGGCGGCGGCGGCGGCTCCTGGTATGCCAAGG
RAG-2(352-369A_{18})	Divergent PCR	F: GCCGCTGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGAAGATCCTGGGGACTC R: GCCGCTGCAGCGGCGGCGGCGGCGGCGGCGGCAGAGCATCTCAAAGTATAG
RAG-2(370-387A_{18})	Divergent PCR	F: GCCGCTGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAAGCAACCAGTTTTG R: GCCGCTGCAGCGGCGGCGGCGGCGGCGGCGGCTGTTGATGTCTGACTGTTG
RAG-2(388-396A_{18})	Divergent PCR	F: GCCGCTGCAGCTGCTGCTGCTGATGAATTTGACACCTAC R: GCCGCTGCAGCGGCGGCGGCACTGAAACAAAATTCCTC
RAG-2(397-405A_{18})	Divergent PCR	F: GCCGCTGCAGCTGCTGCTGCTGATGAAGATGACGAGTC R: GCCGCTGCAGCGGCGGCGGCGTCACCATCAAAACTGG

RAG-2(388-405A$_{18}$)	Divergent PCR	F: GCCGCTGCAGCTGCTGCTGCTGCTGCTGCTGC TGCTGATGAAGATGACGAGTC R: GCCGCTGCAGCGGCGGCGGCGGCGGCGGCGGC ACTGAAACAAAATTCCTC
RAG-2(406-414A$_{18}$)	Divergent PCR	F: GCCGCTGCAGCTGCTGCTGCTTACTGGATAAC ATGTTGC R: GCCGCTGCAGCGGCGGCGGCATCTTCATTGTA GGTGTC
RAG-2(415-424A$_{18}$)	Divergent PCR	F: GCCGCTGCAGCTGCTGCTGCTGTTGACATCAA TACCTGG R: GCCGCTGCAGCGGCGGCGGCGCCGGTTACAGA CTCGTC
RAG-2(D/E370-405A)	GeneArt Strings	GGCGCCCCCTGGGTACCCCAGCAGTGAATTGCACA GTCTTGCCAGGAGGAATCTCTGTCTCCAGTGCAAT CCTCACTCAAACAAACAATGATGAATTTGTTATTG TGGGTGGTTATCAGCTGGAAAATCAGAAAAGGATG GTCTGCAGCCTTGTCTCTCTAGGGGACAACACGAT TGAAATCAGTGAGATGGAGACTCCTGACTGGACCT CAGATATTAAGCATAGCAAAATATGGTTTGGAAGC AACATGGGAAACGGGACTATTTTCCTTGGCATACC AGGAGACAATAAGCAGGCTATGTCAGAAGCATTCT ATTTCTATACTTTGAGATGCTCTGAAGAGGATTTG AGTGAAGATCAGAAAATTGTCTCCAACAGTCAGAC ATCAACAGCAGCTCCTGGTGCCTCCACTCCCTTTG CAGCCTCAGCCGCATTTTGTTTCAGTGCTGCAGCA ACCAGTTTTGCTGGTGCCGCTGCATTTGCCACCTA CAATGCAGCTGATGAAGATGACGAGTCTGTAACCG GCTACTGGATAACATGTTGCCCTACTTGTGATGTT GACATCAATACCTGGGTTCCGTTCTATTCAACGGA GCTCAATAAACCCGCCATGATCTATTGTTCTCATG GGGATGGGCACTGGGTACATGCCCAGTGCATGGGG CGCC
RAG-2(P370-405A)	GeneArt Strings	CCGCGGCCCTGGGTACCCCAGCAGTGAATTGCACA GTCTTGCCAGGAGGAATCTCTGTCTCCAGTGCAAT CCTCACTCAAACAAACAATGATGAATTTGTTATTG TGGGTGGTTATCAGCTGGAAAATCAGAAAAGGATG GTCTGCAGCCTTGTCTCTCTAGGGGACAACACGAT TGAAATCAGTGAGATGGAGACTCCTGACTGGACCT CAGATATTAAGCATAGCAAAATATGGTTTGGAAGC AACATGGGAAACGGGACTATTTTCCTTGGCATACC AGGAGACAATAAGCAGGCTATGTCAGAAGCATTCT ATTTCTATACTTTGAGATGCTCTGAAGAGGATTTG AGTGAAGATCAGAAAATTGTCTCCAACAGTCAGAC ATCAACAGAAGATCCTGGTGACTCCACTCCCGCTG AAGACTCAGAGGAAGCTTGTGCCAGTGCTGAAGCA ACCAGTGCTGATGGTGACGATGAAGCTGACACCTA CAATGAAGATGATGAAGATGACGAGTCTGTAACCG

		GCTACTGGATAACATGTTGCCCTACTTGTGATGTT GACATCAATACCTGGGTTCCGTTCTATTCAACGGA GCTCAATAAACCCGCCATGATCTATTGTTCTCATG GGGATGGGCACTGGGTACATGCCCAGTGCATGGGG CGCC
RAG-2(S/T370-450A)	GeneArt Strings	CCGCGGCCCTGGGTACCCCAGCAGTGAATTGCACA GTCTTGCCAGGAGGAATCTCTGTCTCCAGTGCAAT CCTCACTCAAACAAACAATGATGAATTTGTTATTG TGGGTGGTTATCAGCTGGAAAATCAGAAAAGGATG GTCTGCAGCCTTGTCTCTCTAGGGGACAACACGAT TGAAATCAGTGAGATGGAGACTCCTGACTGGACCT CAGATATTAAGCATAGCAAAATATGGTTTGGAAGC AACATGGGAAACGGGACTATTTTCCTTGGCATACC AGGAGACAATAAGCAGGCTATGTCAGAAGCATTCT ATTTCTATACTTTGAGATGCTCTGAAGAGGATTTG AGTGAAGATCAGAAAATTGTCTCCAACAGTCAGAC ATCAACAGAAGATCCTGGTGACGCCGCTCCCTTTG AAGACGCAGAGGAATTTTGTTTCGCTGCTGAAGCA GCCGCTTTTGATGGTGACGATGAATTTGACGCCTA CAATGAAGATGATGAAGATGACGAGTCTGTAACCG GCTACTGGATAACATGTTGCCCTACTTGTGATGTT GACATCAATACCTGGGTTCCGTTCTATTCAACGGA GCTCAATAAACCCGCCATGATCTATTGTTCTCATG GGGATGGGCACTGGGTACATGCCCAGTGCATGGCC GCGG
RAG-2(D/E352-405A)	GeneArt Strings	GGCGCCCCCTGGGTACCCCAGCAGTGAATTGCACA GTCTTGCCAGGAGGAATCTCTGTCTCCAGTGCAAT CCTCACTCAAACAAACAATGATGAATTTGTTATTG TGGGTGGTTATCAGCTGGAAAATCAGAAAAGGATG GTCTGCAGCCTTGTCTCTCTAGGGGACAACACGAT TGAAATCAGTGAGATGGAGACTCCTGACTGGACCT CAGATATTAAGCATAGCAAAATATGGTTTGGAAGC AACATGGGAAACGGGACTATTTTCCTTGGCATACC AGGAGACAATAAGCAGGCTATGTCAGAAGCATTCT ATTTCTATACTTTGAGATGCTCTGCGGCCGCTTTG AGTGCAGCTCAGAAAATTGTCTCCAACAGTCAGAC ATCAACAGCAGCTCCTGGTGCCTCCACTCCCTTTG CAGCCTCAGCCGCATTTTGTTTCAGTGCTGCCGCA ACCAGTTTTGCTGGTGCCGCTGCATTTGCCACCTA CAATGCAGCTGATGAAGATGACGAGTCTGTAACCG GCTACTGGATAACATGTTGCCCTACTTGTGATGTT GACATCAATACCTGGGTTCCGTTCTATTCAACGGA GCTCAATAAACCCGCCATGATCTATTGTTCTCATG GGGATGGGCACTGGGTACATGCCCAGTGCATGGGG CGCC
pCLIP2A	PCR	F: GCGATCGATACCATGTCCCTGCAGATGGTAAC R: GCGCTCGAGCGGCCGCTTATAACGCGCGC

RAG-2 truncation F	PCR	F: GCTCGGATCCCGGGTACC
RAG-2(1-351) R	PCR	R: GAATTCAGAGCATCTCAAAGTATAGAAATA
RAG-2(1-387) R	PCR	R: GAATTCACTGAAACAAAATTCCTCTGAG
RAG-2(1-405) R	PCR	R: GAATTCATCTTCATTGTAGGTGTCAAATTC
RAG-2(1-405, D/E352-405A) R	PCR	R: GAATTCAGCTGCATTGTAGGTGGCAAA

Assays for exogenous rearrangement

Extrachromosomal rearrangement assays were performed with 10 μg each of MBP-RAG-1-myc-his (full length) and MBP-RAG-2-myc-his (wild-type or variant, as indicated), and 4 μg either pJH200 or pJH290, described previously(Hesse et al. 1987). Plasmids were transfected into NIH3T3 cells in a 10 cm dish with TransIT-LT1 (Mirus, Cat#MIR 2300) or Lipofectamine 20000 (Thermo Fisher, Cat#11668-019) per manufacturer's protocol. Recombination efficiency was similar with both protocols (results not shown). After 48 hours, the cell pellet was divided into thirds, one for storage, one for DNA isolation, and one for protein extraction. Plasmid DNA was extracted by a modified Hirt extraction (Qiagen, Cat#27104) per manufacturer's protocol. DNA (3 μL, about 40 μg) was transformed into 50 μL DH5α Max Efficiency cells (Thermo Fisher, Cat#18258012) per manufacturer's protocol. 1.7% of the transformation mixture was plated on LB agar containing 50 μg / mL ampicillin and the rest was plated on LB agar containing 50 μg / mL ampicillin and 20 μg / mL chloramphenicol. Plates containing ampicillin alone were scored after 16 hours at 37 °C, while plates with ampicillin and chloramphenicol were scored at 20 hours. Protein extraction was performed by addition of 150 μL of boiling SDS lysis buffer [60 mM Tris (pH 7.6), 1% SDS] and boiling for 30 min.

Immunoblotting

For western blots, 60-90 μg protein (calculated using BioRad DC Protein Assay, Cat#500-0116) was loaded on an 8% SDS-PAGE gel and transferred to a 0.45 μm nitrocellulose membrane (Bio-Rad, Cat#1620115). Membranes were blocked at least 45 min with 5% milk in PBST [1X PBS, 0.1% Tween].

Assays for endogenous recombination

The retroviral vector pCLIP2A (Pomerantz et al. 2002) was programmed to coexpress puromycin N-acetyl transferase and RAG-2 variants. The RAG-2 cassettes were amplified from the corresponding pcDNA3.1 subclones by PCR using primers indicated in Table 1. Viral particles were generated by cotransfection of pCLIP2A constructs and pCL-Eco into 293T cells and concentrated by centrifugation. The B progenitor cell lines 63-12 and R2K3 were infected by spin inoculation in the presence of 10 μg/mL polybrene. Cells were maintained under selection with 1 μg/mL puromycin for 21 - 25 days. Recombination was induced in R2K3 cells (10 ml at 10^6 cells/ml) by addition of STI-571 (3 mM stock in DMSO) to a concentration of 3 μM for 48 or 96 hr.
Genomic DNA was isolated from approximately 10^6 cells (Qiagen DNeasy) and D_H-to-J_H or Vκ-to-Jκ rearrangements were detected by PCR. Rearrangements were amplified from 100 ng genomic DNA template by PCR using primers and annealing temperatures indicated in Table 2. RAG-1 was amplified for purposes of normalization from 30 ng genomic DNA template. Reaction products were fractionated by electrophoresis through 1% agarose, transferred to Hybond N+ membranes (GE Healthcare) and detected by hybridization to radiolabeled probes defined in Table 2.

Table 2. Primers and probes used to assay rearrangement in B-cell progenitors

Primer	Anneal/Hybridization (°C)	Sequence 5' to 3'
DSP2 F	65 Anneal	ATGGCCCCTGACACTCTGCACTGCT
DFL16.1 F	65 Anneal	ACACCTGCAAAACCAGAGACCATA
Jh4 R	65 Anneal	AAAGACCTGGAGAGGCCATTCTTACC
Jh probe	61 Hybridization	CTTACCTGAGGAGACGGTGAC
IgKv F	65 Anneal	GSTTCASWGGCAGTGGRTCTGG
Mar35 R	65 Anneal	AACACTGGATAAAGCAGTTTATGCCCTTTC
IgK probe	58 Hybridization	GCTCATTATCAGTTGACGTGGC
RAG1 F	Anneal	GCATCTATTCTGTAGGATCTGC
RAG1 R	Anneal	AAACAATGTCAAGCAGACAGCC
Imu F	60 Anneal	AATACCCGAAGCATTTACAGTGACT
Imu R	60 Anneal	AAGATTTGTGAAGCCGTTTTGACCA
Mu0 F	60 Anneal	GTGCAGGTCCCTCTCTTGTT
Mu0 R	60 Anneal	GACATTGCATCCACCCTTCT
ACTB F	61 Anneal	TGACGTGGACATCCGCAAAG
ACTB R	61 Anneal	CTGGAAGGTGGACAGCGAGG

Assays for germline transcription

RNA was isolated from R2K3 cells at 48 hours after arrest by STI-571 and cDNA was synthesized by random hexamer priming from total RNA. Sequences corresponding to Iμ or μ0 transcripts were detected by PCR using primers indicated in Table 2. Amplification of cDNA for actin beta (ACTB) was performed as a control (Liu et al. 2015).

Chromatin immunoprecipitation

Transduced B progenitor cells were treated with 3 μM STI571 for 21 hr. Chromatin immunoprecipitation (ChIP) was performed as described for RAG-1 and RAG-2 (Ji et al. 2010) and for H3K4me3 (Chakraborty et al. 2009). Input and immunoprecipitated DNA were quantified performed by PicoGreen staining (Thermo Fisher). Each ChIP was performed in duplicate and each real-time PCR reaction was performed in duplicate. For

analysis of H3K4me3 enrichment 200 pg of DNA was used. The relative abundance of amplicons in the immunoprecipitated DNA relative to input was analyzed by real-time PCR using the primers listed in Table 3. The enrichment (IP/Input-$_{corr}$) of RAG-1 or RAG-2 at specific regions was calculated as described (Ji et al. 2010).

Table 3. Primers used in ChIP analyses

Primer	Sequence 5'-3'	Source
γ-actin FP	GACACCCAACCCCGTGACG	(Subrahmanyam et al. 2012)
γ-actin RP	GCGGCCATCACATCCCAG	
Cγ3 FP	TGGACAAACAGAAGTAGACATGGGTC	
Cγ3RP	GGGGTTTAGAGGAGAGAAGGCAC	
DSP2s FP	TGTTACCTTACTTGGCAGGGATTT	
DSP2s RP	TGGGTTTTTGTTGCTGGATATATC	
DFL16.1 FP	CAAAGCAGCCACCATCCAG	(Chakraborty et al. 2009)
DFL16.1 RP	GCAGCACGGTTGAGTTTCAG	
DQ52 FP	CCCTGTGGTCTCTGACTGGTG	
DQ52 RP	GATTTCTCAAGCCTCTCTACTTCCTC	
JH2 FP	TACTTTGACTACTGGGGC	
JH2 RP	CCCTAGTCCTTCATGACC	
Jκ1 FP	TTGTACAGCCAGACAGTGGAG	(Ji et al. 2010)
Jκ1 RP	GCCACAGACATAGACAACGG	
Jκ2 FP	CAGATTCTGGCACTCTCCAA	
Jκ2 RP	ACTGAGCATGGTCTGAGCAC	
Jκ4 FP	AGTGTGAAAGCTGAGCGAAA	
Jκ4 RP	CACAGTGAGGACTATGACATGC	
Cκ FP	GTGGAAGATTGATGGCAGTG	
Cκ RP	GCTCATGCTGTAGGTGCTGT	
Fahd1 FP	AGAGACCTTTCGCTGACCTC	
Fahd1 RP	GGTCATGTGACCACCGACT	
Rik FP	GTTTCCACCGGAAGTGCT	
Rik RP	GCGCGCTAATAGGGTCTTT	

Oligonucleotide substrates

The duplex oligonucleotides used in the biochemical assays were HL44/45 (12-RSS) and HL46/47 (23-RSS) (Shimazaki et al. 2012). Oligonucleotides were purified by gel

electrophoresis and extraction and end-labeled where indicated with 32P by T4 DNA polynucleotide kinase (New England Biolabs).

Table 4. Oligonucleotide substrates for biochemical assays

Oligo	Sequence 5'-3'
HL44	GGGCTATACGCAGCTTGGCTGCAGGTCGACCACAGTGCTACAGACTGGAACA AAAACCCTGCAGTCGA
HL45	TCGACTGCAGGGTTTTTGTTCCAGTCTGTAGCACTGTGGTCGACCTGCAGCC AAGCTGCGTATAGCCC
HL46	GGGTCTCGCCAAGCTGATCCCCGGGGATCCCACAGTGGTAGTACTCCACTGT CTGGCTGTACAAAAACCCTCGGGATCC
HL47	GGATCCCGAGGGTTTTTGTACAGCCAGACAGTGGAGTACTACCACTGTGGGA TCCCCGGGGATCAGCTTGGCGAGACCC

Protein purification

Core RAG-1 (cR1-MH), the core RAG-1ct variant (cR1ct-MH) or full-length RAG-1 (flR1- MH) were co-expressed in HEK 293T cells with full-length RAG-2 (fR2-MH) or the corresponding RAG-2 mutants. RAG complexes were purified as described (Raval et al. 2008). Transfected 293T cells were harvested in PBS-EDTA [1X PBS, 2 mM EDTA] and pelleted at 500 g at 4°C for 5 minutes. The supernatant was aspirated and cell pellets were resuspended in 7ml buffer R [25 mM HEPES (pH 7.4), 150 mM KCl, 2 mM DTT, 1.04 mM aminoethyl-benzenesulfonyl fluoride, 0.8 μM aprotinin, 40 nM bestatin, 14 nM E-64, 20 nM leupeptin, 15 nM pepstatin A and 10% glycerol]. Cell suspensions were placed in a 50% ethanol dry ice bath and subjected to 3 rounds of sonication in a Branson Digital Sonifier 450. Each round of sonication was performed at 23% amplitude for 1.5 12 minutes, with 30 second intervals of sonication followed by a 10 second rest between intervals. The lysate was clarified by centrifugation at 46,000 g in an SW55Ti rotor at

4°C for 30 min and the supernatant was loaded onto 1 ml amylose resin that had been equilibrated with buffer R. The column was washed once with 5 ml buffer R and once with 5 ml buffer R lacking protease inhibitors. Protein was eluted with 10 mM maltose in buffer R lacking protease inhibitors and then dialyzed against buffer R. Aliquots were snap frozen and stored at -80˚C.

Burst kinetic analysis

Varying nominal concentrations of RAG protein were combined with 200 nM total radiolabeled HL44/45 in binding buffer containing 1% glycerol (reaction volume 10 µl) and incubated for 20 min at 37°C, at which time $MgCl_2$ was added to 5 mM. Incubation was continued for an additional 30 min at 37°C. Reactions were stopped by addition of 10 µl 90% formamide-TBE and heated for 5 min at 95°C. Products were fractioned by electrophoresis on a 15% polyacrylamide-urea gel, visualized by a phosphorimager quantified using ImageQuantNL. The active fraction of each RAG preparation was determined by (1) plotting accumulation of nicked product ([N]) as a function of time for each concentration of total RAG; (2) extrapolating rates to zero time, thereby obtaining the initial burst of nicked product formation ($[N]_0$) at each nominal RAG concentration; and (3) expressing $[N]_0$ as a linear function of nominal RAG concentration:

$$[N]_0 = f_a \bullet [RAG]_T + b$$

where $[RAG]_T$ is the total (nominal) concentration of RAG, assuming a tetrameric stoichiometry of $(RAG\text{-}1)_2(RAG\text{-}2)_2$, and fa is the fraction of $[RAG]_T$ that is active (Yu & Lieber 2000).

Assays for coupled cleavage

RAG (1 nM active tetramer, as determined by burst kinetic analysis) was combined in binding buffer with 5 nM HL44/45, 5 nM HL46/47 and varying concentrations of a peptide corresponding to the amino-terminal 21 residues of histone H3, either trimethylated at lysine 4 (H3K4me3; Anaspec, 64194) or unmethylated (H3K4me0; Anaspec, 61701), in a reaction volume of 10 µl. After incubation for 20 min at 37°C, $MgCl_2$ was added to 5 mM and incubation was continued for an additional 1 hr. Reactions were stopped and products were analyzed as above.

Assays for DNA nicking

RAG (1 nM active tetramer, as determined by burst kinetic analysis) was combined in binding buffer with 5 nM HL44/45 and varying concentrations of a peptide corresponding to the amino-terminal 21 residues of histone H3, either trimethylated at lysine 4 (H3K4me3; Anaspec, 64194) or unmethylated (H3K4me0; Anaspec, 61701), in a reaction volume of 10 µl. After incubation for 20 min at 37°C, $MgCl_2$ was added to 5 mM and incubation was continued for an additional 30 min. Reactions were stopped and products were analyzed as above.

References

Akamatsu, Y. et al., 2003. Deletion of the RAG2 C terminus leads to impaired lymphoid development in mice. *Proceedings of the National Academy of Sciences of the United States of America*, 100(3), pp.1209–14.

Akira, S., Okazaki, K. & Sakano, H., 1987. Two Pairs of Recombination Signals are Sufficient to Cause Immunoglobulin V-(D)-J Joining. *Science*, 238(4830), pp.1134–1138.

Alessandrini, A. & Desiderio, S. V, 1991. Coordination of immunoglobulin DJH transcription and D-to-JH rearrangement by promoter-enhancer approximation. *Molecular and cellular biology*, 11(4), pp.2096–107.

Alt, F.W. et al., 1984. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *The EMBO journal*, 3(6), pp.1209–19.

Aoki-Ota, M. et al., 2012. Skewed Primary Igk Repertoire and V–J Joining in C57BL/6 Mice: Implications for Recombination Accessibility and Receptor Editing. *The Journal of Immunology*, 188(5), pp.2305–2315.

Askary, A. et al., 2014. Modeling of the RAG Reaction Mechanism. *Cell Reports*, 7(2), pp.307–315.

Bassing, C.H. et al., 2000. Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. *Nature*, 405(6786), pp.583–586.

Besmer, E. et al., 1998. Hairpin Coding End Opening Is Mediated by RAG1 and RAG2 Proteins. *Molecular Cell*, 2(6), pp.817–828.

Bettridge, J. et al., 2017. H3K4me3 induces allosteric conformational changes in the DNA-binding and catalytic regions of the V(D)J recombinase. *Proceedings of the*

National Academy of Sciences of the United States of America, 114(8), pp.1904–1909.

Bolland, D.J. et al., 2007. Antisense intergenic transcription precedes Igh D-to-J recombination and is controlled by the intronic enhancer Emu. *Molecular and cellular biology*, 27(15), pp.5523–33.

Callebaut, I. & Mornon, J.-P., 1998. The V(D)J recombination activating protein RAG2 consists of a six-bladed propeller and a PHD fingerlike domain, as revealed by sequence analysis. *Cellular and Molecular Life Sciences (CMLS)*, 54(8), pp.880–891.

Chakraborty, T. et al., 2009. A 220-nucleotide deletion of the intronic enhancer reveals an epigenetic hierarchy in immunoglobulin heavy chain locus activation. *Journal of Experimental Medicine*, 206(5), pp.1019–1027.

Chakraborty, T. et al., 2007. Repeat Organization and Epigenetic Regulation of the DH-Cμ Domain of the Immunoglobulin Heavy-Chain Gene Locus. *Molecular Cell*, 27(5), pp.842–850.

Chaumeil, J. et al., 2013. The RAG2 C-terminus and ATM protect genome integrity by controlling antigen receptor gene cleavage. *Nature Communications*, 4, p.2231.

Choi, N.M. & Feeney, A.J., 2014. CTCF and ncRNA Regulate the Three-Dimensional Structure of Antigen Receptor Loci to Facilitate V(D)J Recombination. *Frontiers in immunology*, 5, p.49.

Coleclough, C. et al., 1981. Abberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression. *Nature*, 290(5805), pp.372–8.

Corneo, B. et al., 2007. Rag mutations reveal robust alternative end joining. *Nature*,

449(7161), pp.483–6.

Cortes, P., Ye, Z.-S. & Baltimore, D., 1994. RAG-1 Interacts with the Repeated Amino Acid Motif of the Human Homologue of the Yeast Protein SRP1. *Proceedings of the National Academy of Sciences of the United States of America Cell Biology*, 91(16), pp.7633–7637.

Coussens, M.A. et al., 2013. RAG2's Acidic Hinge Restricts Repair-Pathway Choice and Promotes Genomic Stability. *Cell Reports*, 4(5), pp.870–878.

Cuomo, C.A. & Oettinger, M.A., 1994. Analysis of regions of RAG-2 important for V(D)J recombination. *Nucleic Acids Research*, 22(10), pp.1810–1814.

Deng, Z., Liu, H. & Liu, X., 2015. RAG1-mediated ubiquitylation of histone H3 is required for chromosomal V(D)J recombination. *Cell Research*, 25(2), pp.181–192.

Deriano, L. & Roth, D.B., 2013. Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage. *Annual Review of Genetics*, 47, pp.433–455.

Du, H. et al., 2008. Activation of 12/23-RSS-Dependent RAG Cleavage by hSWI/SNF Complex in the Absence of Transcription. *Molecular Cell*, 31(5), pp.641–649.

Elkin, S.K. et al., 2005. A PHD finger motif in the C terminus of RAG2 modulates recombination activity. *The Journal of biological chemistry*, 280(31), pp.28701–10.

Fang Yin, F. et al., 2009. Structure of the RAG1 nonamer binding domain with DNA reveals a dimer that mediates DNA synapsis. *2Nature Structural & Molecular Biology*, 16(5), pp.449–508.

Fisher, M.R. et al., 2017. Immature Lymphocytes Inhibit Rag1 and Rag2 Transcription and V(D)J Recombination in Response to DNA Double-Strand Breaks. *The Journal*

of Immunology, 198, p.published online. Available at: http://www.jimmunol.org/content/early/2017/02/17/jimmun

Gapud, E.J. et al., 2011. Repair of Chromosomal RAG-Mediated DNA Breaks by Mutant RAG Proteins Lacking Phosphatidylinositol 3-Like Kinase Consensus Phosphorylation Sites. *The Journal of Immunology*, 187(4), pp.1826–1834.

Gapud, E.J. & Sleckman, B.P., 2011. Unique and redundant functions of ATM and DNA-PKcs during V(D)J recombination. *Cell Cycle*, 10(12), pp.1928–1935.

van Gent, D.C. et al., 1997. Stimulation of V(D)J cleavage by high mobility group proteins. *The EMBO journal*, 16(10), pp.2665–70.

Gigi, V. et al., 2014. RAG2 mutants alter DSB repair pathway choice in vivo and illuminate the nature of "alternative NHEJ". *Nucleic acids research*, 42(10), pp.6352–64.

Golding, A. et al., 1999. Nucleosome structure completely inhibits in vitro cleavage by the V(D)J recombinase. *The EMBO journal*, 18(13), pp.3712–23.

Goldmit, M. et al., 2005. Epigenetic ontogeny of the Igk locus during B cell development. *Nature Immunology*, 6(2), pp.198–203.

Gomez, C.A. et al., 2000. Mutations in conserved regions of the predicted RAG2 kelch repeats block initiation of V(D)J recombination and result in primary immunodeficiencies. *Molecular and cellular biology*, 20(15), pp.5653–64.

Grazini, U. et al., 2010. The RING Domain of RAG1 Ubiquitylates Histone H3: A Novel Activity in Chromatin-Mediated Regulation of V(D)J Joining. *Molecular Cell*, 37(2), pp.282–293.

Grundy, G.J. et al., 2009. Initial Stages of V(D)J Recombination: The Organization of

RAG1/2 and RSS DNA in the Postcleavage Complex. *Molecular Cell*, 35(2), pp.217–227.

Grundy, G.J., Yang, W. & Gellert, M., 2010. Autoinhibition of DNA cleavage mediated by RAG1 and RAG2 is overcome by an epigenetic signal in V(D)J recombination. *Proceedings of the National Academy of Sciences of the United States of America*, 107(52), pp.22487–92.

Helmink, B.A. & Sleckman, B.P., 2012. The Response to and Repair of RAG-Mediated DNA Double-Strand Breaks. *Annual Review of Immunology*, 30(1), pp.175–202.

Hesse, J.E. et al., 1987. Extrachromosomal DNA Substrates in Pre-B Cells Undergo Inversion or Deletion at lmmunoglobulin V-(D) J Joining Signals. *Cell*, 49, pp.775–783.

Inlay, M. et al., 2002. Essential roles of the κ light chain intronic enhancer and 3′ enhancer in κ rearrangement and demethylation. *Nature immunology*, 3(5), pp.463–468.

Ji, Y. et al., 2010. The In Vivo Pattern of Binding of RAG1 and RAG2 to Antigen Receptor Loci. *Cell*, 141(3), pp.419–431.

Jiang, H. et al., 2005. Ubiquitylation of RAG-2 by Skp2-SCF Links Destruction of the V(D)J Recombinase to the Cell Cycle. *Molecular Cell*, 18(6), pp.699–709.

Johnston, C.M. et al., 2006. Complete Sequence Assembly and Characterization of the C57BL/6 Mouse Ig Heavy Chain V Region. *The Journal of Immunology*, 176(7), pp.4221–4234.

Jones, J.M. et al., 2011. The RAG1 V(D)J recombinase/ubiquitin ligase promotes ubiquitylation of acetylated, phosphorylated histone 3.3. *Immunology Letters*,

136(2), pp.156–162.

Jones, J.M. & Gellert, M., 2003. Autoubiquitylation of the V(D)J Recombinase Protein RAG1 Autoubiquitylation of the V(D)J recombinase protein RAG1. *Source: Proceedings of the National Academy of Sciences of the United States of America*, 100(26), pp.15446–15451.

Kassmeier, M.D. et al., 2012. VprBP binds full-length RAG1 and is required for B-cell development and V(D)J recombination fidelity. *The EMBO journal*, 31(4), pp.945–58.

Kim, D.R. et al., 1999. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. *Genes & development*, 13(23), pp.3070–80.

Kim, M.S. et al., 2015. Crystal structure of the V(D)J recombinase RAG1–RAG2. *Nature*, 518(7540), pp.507–511.

Kirch, S.A., Sudarsanam, P. & Oettinger, M.A., 1996. Regions of RAG1 protein critical for V(D)J recombination. *European Journal of Immunology*, 26(4), pp.886–891.

Kosak, S.T. et al., 2002. Subnuclear Compartmentalization of Immunoglobulin Loci During Lymphocyte Development. *Science*, 296(5565), pp.158–62.

Kwon, J. et al., 2000. Histone Acetylation and hSWI/SNF Remodeling Act in Concert to Stimulate V(D)J Cleavage of Nucleosomal DNA. *Molecular Cell*, 6(5), pp.1037–1048.

Laney, J.D. & Hochstrasser, M., 1999. Substrate Targeting in the Ubiquitin System. *Cell*, 97(4), pp.427–430.

Lee, J. & Desiderio, S., 1999. Cyclin A/CDK2 Regulates V(D)J Recombination by Coordinating RAG-2 Accumulation and DNA Repair. *Immunity*, 11(6), pp.771–781.

Lennon, G.G. & Perry, R.P., 1985. Cmu-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5'-nontranslatable exon. *Nature*, 318(5), pp.475–8.

Lescale, C. et al., 2016. RAG2 and XLF/Cernunnos interplay reveals a novel role for the RAG complex in DNA repair. *Nature Communications*, 7, p.10529.

Li, Z. et al., 1996. A Conserved Degradation Signal Regulates RAG-2 Accumulation during Cell Division and Links V(D)J Recombination to the Cell Cycle. *Immunity*, 5(6), pp.575–589.

Liang, H.-E. et al., 2002. The "Dispensable" Portion of RAG2 Is Necessary for Efficient V-to-DJ Rearrangement during B and T Cell Development. *Immunity*, 17(5), pp.639–651.

Lin, S.G. et al., 2015. CTCF-binding elements 1 and 2 in the Igh intergenic control region cooperatively regulate V(D)J recombination. *Proceedings of the National Academy of Sciences of the United States of America*, 112(6), pp.1815–20.

Lin, W.-C. & Desiderio, S., 1994. Cell cycle regulation of V(D)J recombination-activating protein RAG-2. *Proceedings of the National Academy of Sciences of the United States of America Biochemistry*, 91(7), pp.2733–2737.

Lin, W.-C. & Desiderio, S., 1993. Regulation of V(D)J Recombination Activator Protein RAG-2 by Phosphorylation. *Science*, 260(5110), pp.953–959.

Liu, L.-L. et al., 2015. Identification of Valid Reference Genes for the Normalization of RT-qPCR Expression Studies in Human Breast Cancer Cell Lines Treated with and without Transient Transfection. *PloS one*, 10(1), p.e0117058.

Liu, Y. et al., 2007. A Plant Homeodomain in Rag-2 that Binds Hypermethylated Lysine

4 of Histone H3 Is Necessary for Efficient Antigen-Receptor-Gene Rearrangement. *Immunity*, 27(4), pp.561–571.

Liu, Y., 2009. *Epigenetic control of V(D)J recombination*.

Lu, C. et al., 2015. An Autoregulatory Mechanism Imposes Allosteric Control on the V(D)J Recombinase by Histone H3 Methylation. *Cell Reports*, 10(1), pp.29–38.

Malkova, A. & Ira, G., 2013. Break-induced replication: functions and molecular mechanism. *Current Opinion in Genetics & Development*, 23(3), pp.271–279.

Maman, Y. et al., 2016. RAG1 targeting in the genome is dominated by chromatin interactions mediated by the non-core regions of RAG1 and RAG2. *Nucleic acids research*, 44(20), pp.9624–9637.

Matthews, A.G.W. et al., 2007. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature*, 450(7172), pp.1106–10.

McBlane, F. & Boyes, J., 2000. Stimulation of V(D)J recombination by histone acetylation. *Current Biology*, 10(8), pp.483–486.

McBlane, J.F. et al., 1995. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell*, 83(3), pp.387–395.

McMurry, M.T. & Krangel, M.S., 2000. A role for histone acetylation in the developmental regulation of VDJ recombination. *Science*, 287(5452), pp.495–8.

Medvedovic, J. et al., 2013. Flexible Long-Range Loops in the VH Gene Region of the Igh Locus Facilitate the Generation of a Diverse Antibody Repertoire. *Immunity*, 39(2), pp.229–244.

Morshead, K.B. et al., 2003. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at

lysine 4. *Proceedings of the National Academy of Sciences of the United States of America*, 100(20), pp.11577–82.

Mostoslavsky, R. et al., 1998. kappa chain monoallelic demethylation and the establishment of allelic exclusion. *Genes & Development*, 12(12), pp.1801–1811.

Van Ness, B.G. et al., 1981. Transcription of the unrearranged mouse Cκ locus: Sequence of the initiation region and comparison of activity with a rearranged Vκ-Cκ gene. *Cell*, 27(3), pp.593–602.

Pannunzio, N.R. et al., 2014. Non-homologous end joining often uses microhomology: Implications for alternative end joining. *DNA Repair*, 17, pp.74–80.

Pomerantz, J.L., Denny, E.M. & Baltimore, D., 2002. CARD11 mediates factor-specific activation of NF-[kappa]B by the T cell receptor complex. *EMBO Journal*, 21(19), pp.5184–5194.

Ramón-Maiques, S. et al., 2007. The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. *Proceedings of the National Academy of Sciences of the United States of America*, 104(48), pp.18993–8.

Ramsden, D.A. et al., 1996. Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. *The EMBO Journal*, 15(12), pp.3197–3206.

Raval, P. et al., 2008. Evidence for Ku70/Ku80 association with full-length RAG1. *Nucleic Acids Research*, 36(6), pp.2060–2072.

Ross, A.E., Vuica, M. & Desiderio, S., 2003. Overlapping signals for protein degradation and nuclear localization define a role for intrinsic RAG-2 nuclear uptake in dividing cells. *Molecular and cellular biology*, 23(15), pp.5308–19.

Ru, H. et al., 2015. Molecular Mechanism of V(D)J Recombination from Synaptic RAG1-RAG2 Complex Structures. *Cell*, 163(5), pp.1138–1152.

Sadofsky, M.J. et al., 1993. Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic acids research*, 21(24), pp.5644–50.

Sadofsky, M.J., Hesse, J.E. & Gellert, M., 1994. Definition of a core region of RAG-2 that is functional in V(D)J recombination. *Nucleic Acids Research*, 22(10), pp.1805–1809.

Sakano, H. et al., 1979. Sequences at the somatic recombination site of immunoglubulin light-chain genes. *Nature*, 280, pp.288–294.

Sekiguchi, J., Whitlow, S. & Alt, F.W., 2001. Increased Accumulation of Hybrid V(D)J Joins in Cells Expressing Truncated versus Full-Length RAGs. *Molecular Cell*, 8(6), pp.1383–1390.

Selimyan, R. et al., 2013. Localized DNA demethylation at recombination intermediates during immunoglobulin heavy chain gene assembly. D. Nemazee, ed. *PLoS biology*, 11(1), p.e1001475.

Shimazaki, N. et al., 2012. Mechanistic basis for RAG discrimination between recombination sites and the off-target sites of human lymphomas. *Molecular and cellular biology*, 32(2), pp.365–75.

Shimazaki, N., Tsai, A.G. & Lieber, M.R., 2009. H3K4me3 Stimulates the V(D)J RAG Complex for Both Nicking and Hairpinning in trans in Addition to Tethering in cis: Implications for Translocations. *Molecular Cell*, 34(5), pp.535–544.

Shockett, P.E. & Schatz, D.G., 1999. DNA hairpin opening mediated by the RAG1 and RAG2 proteins. *Molecular and cellular biology*, 19(6), pp.4159–66.

Silver, D.P. et al., 1993. Dispensable sequence motifs in the RAG-I and RAG-2 genes for plasmid V(D)J recombination. *Immunology*, 90, pp.6100–6104.

Simkus, C. et al., 2007. Biochemical and Folding Defects in a RAG1 Variant Associated with Omenn Syndrome. *The Journal of Immunology*, 179, pp.8332–8340.

Simkus, C., Bhattacharyya, A., et al., 2009. Correlation between recombinase activating gene 1 ubiquitin ligase activity and V(D)J recombination. *Immunology*, 128(2), pp.206–17.

Simkus, C., Makiya, M. & Jones, J.M., 2009. Karyopherin alpha 1 is a putative substrate of the RAG1 ubiquitin ligase. *Molecular Immunology*, 46(7), pp.1319–1325.

Singh, S.K. & Gellert, M., 2015. Role of RAG1 autoubiquitination in V(D)J recombination. *Proceedings of the National Academy of Sciences of the United States of America*, 112(28), pp.8579–83.

Stanhope-Baker, P. et al., 1996. Cell Type–Specific Chromatin Structure Determines the Targeting of V(D)J Recombinase Activity In Vitro. *Cell*, 85, pp.887–897.

Steen, S.B. et al., 1999. Roles of the "dispensable" portions of RAG-1 and RAG-2 in V(D)J recombination. *Molecular and cellular biology*, 19(4), pp.3010–3017.

Su, L.-K. & Kadesch, T., 1990. The immunoglobulin heavy-chain enhancer functions as the promoter for Imu sterile transcription. *Molecular and cellular biology*, 10(6), pp.2619–24.

Subrahmanyam, R. et al., 2012. Localized epigenetic changes induced by DH recombination restricts recombinase to DJH junctions. *Nature Immunology*, 13(12), pp.1205–1212.

Swanson, P.C., 2001. The DDE motif in RAG-1 is contributed in trans to a single active

site that catalyzes the nicking and transesterification steps of V(D)J recombination. *Molecular and cellular biology*, 21(2), pp.449–58.

Villa, A. et al., 2001. V(D)J recombination defects in lymphocytes due to RAG mutations: severe immunodeficiency with a spectrum of clinical presentations. *Blood*, 97(1), pp.81–88.

West, K.L. et al., 2005. A direct interaction between the RAG2 C terminus and the core histones is required for efficient V(D)J recombination. *Immunity*, 23(2), pp.203–212.

Xu, C.-R. & Feeney, A.J., 2009. The Epigenetic Profile of Ig Genes Is Dynamically Regulated during B Cell Differentiation and Is Modulated by Pre-B Cell Receptor Signaling. *Journal of immunology*, 182(3), pp.1362–1369.

Yancopoulos, G. & Alt, F.W., 1985. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell*, 40(2), pp.271–281.

Ye, J., 2004. The immunoglobulin IGHD gene locus in C57BL/6 mice. *Immunogenetics*, 56(6), pp.399–404.

Yu, K. & Lieber, M.R., 2000. The nicking step in V(D)J recombination is independent of synapsis: implications for the immune repertoire. *Molecular and cellular biology*, 20(21), pp.7914–21.

Yuan, C.-C. et al., 2012. Histone H3R2 Symmetric Dimethylation and Histone H3K4 Trimethylation Are Tightly Correlated in Eukaryotic Genomes. *Cell Reports*, 1(2), pp.83–90.

Yurchenko, V., Xue, Z. & Sadofsky, M., 2003. The RAG1 N-terminal domain is an E3 ubiquitin ligase. *Genes & development*, 17(5), pp.581–585.

Zagelbaum, J. et al., 2016. Real-time analysis of RAG complex activity in V(D)J

recombination. *Proceedings of the National Academy of Sciences of the United States of America*, 113(42), pp.11853–11858.

Zhang, L. et al., 2011. Coupling of V(D)J Recombination to the Cell Cycle Suppresses Genomic Instability and Lymphoid Tumorigenesis. *Immunity*, 34(2), pp.163–174.

Curriculum Vitae

The Johns Hopkins University School of Medicine

Alyssa Maria Ward **March 7, 2017**

Education:

Ph.D.	2017	Biochemistry, Cellular, and Molecular Biology Johns Hopkins School of Medicine Mentor: Stephen Desiderio, M.D., Ph.D.
B.A.	2009	Molecular Biology Claremont McKenna College

Professional Experience:

Research Assistant	2010-present	Lab of Dr. Stephen Desiderio The Johns Hopkins University School of Medicine Genetically identified a novel inhibitory domain in the RAG recombinase Utilized biochemical and cellular assays for recombination
Manuscript editor	2015-present	Ad hoc English language editing of manuscripts
Intern	2015	Internship in the Office of Public Affairs at the Federation of American Societies for Experimental Biology
Research Assistant	2009-2010	Lab of Dr. Susann Brady-Kalnay Case Western Reserve University Biochemical identification of protein-protein interactions that govern glioblastoma metastasis Utilized microscopy to assay cell motility
Research Assistant	2008-2009	Lab of Dr. Gretchen Edwalds-Gilbert Claremont McKenna College Assessed protein-protein interactions of a splicing factor

Peer Reviewed Publications:

Ward A, Kumari G, Sen R, Desiderio S (submitted - Immunity) The RAG-2 inhibitory domain gates accessibility of the V(D)J recombinase to chromatin.

Lu C, **Ward A**, Bettridge J, Liu Y, and Desiderio S (2015) An Autoregulatory Mechanism Imposes Allosteric Control on the V(D)J Recombinase by Histone H3 Methylation. Cell Reports, 10(1), pp.29–38.

Leung et al. (2015) Drosophila muller f elements maintain a distinct set of genomic properties over 40 million years of evolution. G3, 4;5(5), pp.719-40

Non-Peer Reviewed Publication:

Ward A, Baldwin TO, Antin PB (2016) Research data: Silver lining to irreproducibility. Nature, 532, p.177

Service and leadership:

2015-present	Developed a curriculum for teaching reproducibility and rigor, taught lectures, currently assisting e-lecture development